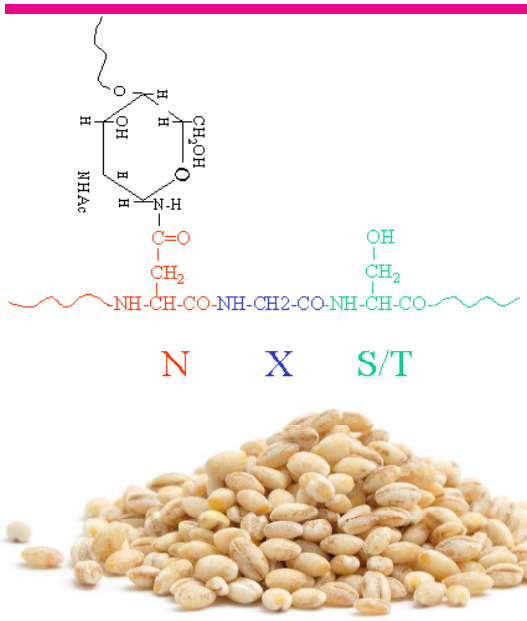
Glycoproteins and Glycosylation Site Assignments in Cereal seed Proteomes

Dedvisitsakul, Plaipol

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Glycoproteins and Glycosylation Site Assignments in Cereal seed Proteomes



Plaipol Dedvisitsakul
Ph.D. Thesis
August 2013

Enzyme and Protein Chemistry
Department of Systems Biology

DTU Systems Biology

Glycoproteins and Glycosylation Site Assignments in Cereal seed Proteomes

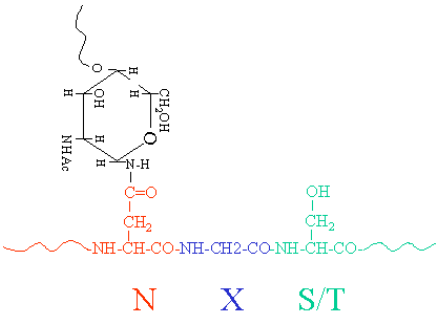
Plaipol Dedvisitsakul

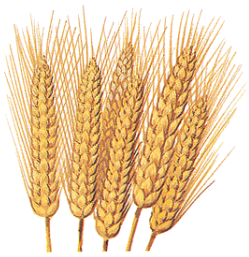
Ph.D. Thesis

August 2013

Enzyme and Protein Chemistry (EPC)

Technical University of Denmark





Supervisors:

Professor Birte Svensson

Associate Professor Per Hägglund

Associate Professor Christine Finnie

Associate Professor Susanne Jacobsen

Enzyme and Protein Chemistry
Department of Systems Biology

Address: Enzyme and Protein Chemistry
Department of Systems Biology
Technical University of Denmark
Søltofts Plads, Building 224
DK-2800 Kgs. Lyngby
Denmark
Phone: +45 4525 2525
Fax: +45 4588 6307
Web: www.bio.dtu.dk

Print: J&R Frydenberg A/S
København
December 2013

ISBN: 978-87-91494-76-5

PREFACE

The present thesis reports the results of my Ph.D. project carried out in the Enzyme and Protein Chemistry group (EPC), Department of Systems Biology, Technical University of Denmark from September 2009 to August 2013 under supervision of Associate Professor Susanne Jacobsen, Associate Professor Per Hägglund, Associate Professor Christine Finnie and Professor Birte Svensson. The Ph.D. stipend was funded by a Thai Government Scholarship.

The work of this Ph.D. project has resulted in the following manuscripts:

Dedvisitsakul, P., Jacobsen, S., Svensson, B., Bunkenborg, J., Finnie, C., Hägglund, P. Selective Glycopeptide Enrichment Using a Combination of ZIC-HILIC and Cotton Wool for Exploring the Subglycoproteome of Wheat Flour Albumins (*Manuscript to be submitted to Journal of Proteome Research).*

Barba-Espín, G., **Dedvisitsakul, P.**, Hägglund, P., Svensson, B., Finnie, C. GA_3-induced aleurone layers responding to heat shock or tunicamycin provide insight into the *N*-glycoproteome, protein secretion and ER stress (*Plant Physiology*).

ACKNOWLEDGEMENTS

I entirely bestow my deepest thanks to my major supervisors Associate Professor Susanne Jacobsen, Associate Professor Per Hägglund, Associate Professor Christine Finnie and Professor Birte Svensson for the opportunity to do this study and their valuable suggestions, active discussions, as well as kind support and contribution in every academic activity I have done throughout these years of my Ph.D. studies. My appreciation extends to past and current members of Enzyme and Protein Chemistry Group for their friendship and support. Special thanks go to Birgit Andersen for her invaluable help with mass spectrometry analysis and training, Louise Helsted and Julie Henriksen for their kindness with preparation of protein samples, Gregorio Barba Espin for his special collaboration, discussions and suggestions as well as Adiphol Dilokpimol for his friendship, support and advices thoughtout the time I have been living in Denmark.

I would like to thank the Royal Thai Government for the financial support and for giving me an invaluable opportunity to pursue my Ph.D. at DTU.

Lastly, I owe my endless love and gratefulness to my parents, Mr. Sanguan and Mrs. Nongyao Dedvisitsakul for their unconditional love, for being the very first persons who educated me, and for always being there for me. I also would like to thank Mrs. Chanitchote Dedvisitsakul, my sister who always loves and supports me. My appreciation also extends to my friends in both in Thailand and Denmark for their consistent encouragement and support.

ABSTRACT

The study of plant proteomes is important to further the understanding of biological processes and enhance the agronomical and nutritional value of crops and food products. To gain deeper understanding on the proteome level, it is important to characterize post-translational modifications. Glycosylation is one of the most common PTMs of protein that is involved in many physiological functions and biological pathways.

The aim of this Ph.D. project is mainly to screen and identify *N*-glycosylated proteins from barley and wheat. A HILIC-based glycopeptide enrichment technqiue was first developed by supplementing cotton wool with ZIC-HILIC in a microcolumn (called ZIC-cotton). This approach reduced co-enrichment of non-glycosylated peptides and allowed glycoppeptide identification from large protein mixtures. It was applied for glycoprotein identification and glycosylation site assignment in wheat albumin and barley aleurone layer proteins. By site-specific glycosylation labeling and LC-MS/MS analysis, 76 different glycosylation sites within 65 wheat albumin proteins were identified using a combination of ZIC-cotton and cotton wool. In addition, ZIC-cotton has been also applied to proteins produced from barley aleurone layer and 47 glycoproteins were identified. Sequence homology search against allergen database reveals that many glycoproteins identified from wheat and barley share similarity with known food allergens and may therefore be targets in search of novel allergens from wheat flour.

DANSK RESUMÉ (SUMMARY IN DANISH)

Proteomanalyse er et vigtigt element i studiet af biologiske processer i planter med henblik på at øge forståelsen af dyrknings- og ernæringsmæssige aspekter af landbrugsafgrøder samt heraf forarbejdede produkter. For at opnå en dybere forståelse af proteomanalyse er det vigtigt at karakterisere post-translationelle modifikationer (PTM) af proteiner. Glykosylering er én af de mest almindelige PTM i proteiner og som er involveret i mange fysiologiske funktioner og biologiske reaktioner.

Udgangspunktet for dette Ph.D.-projekt er at screene og identificere N-glykosylerede proteiner i byg og hvede. Til at berige fraktionen af glykopeptider blev først udviklet hydrofil-interaktionskromatografi (eller hydrofil-interaktionsvæskekromatografi) HILIC-baseret mikrokolonne som er baseret på zwitterion-hydrofil- interaktionsvæskekromatografi) ZIC-HILIC sammen med bomuld (kaldet ZIC-bomuld). Ved kromatografi af ekstrakt af stor proteinblanding og brug af denne ZIC-bomuld-mikrokolonne opnåedes stor berigelse af glykoprotider samt reduktion af co-eluering ikke-glykosylerede protiner. Kromatografi ved brug af mikrokolonnen blev brugt til at identificere glykoproteiner samt glykosyleringssteder i ekstrakter af proteiner fra byg aleuronlag hhv. hvedealbumin. Ved site-specifik glykosyleringsmærkning og LC-MS/MS analyse blev 76 forskellige glykosyleringssteder i 65 hvedealbuminproteiner identificeret ved brug af ZIC-bomuld og bomuld. I bygaleuron-proteinfraktion og brug af ZIC-bomuld blev 49 glykoproteiner identificeret. Sekvenshomologisøgning af identificerede glykoproteiner i allergendatabase viser at mange af de identificerede glykoproteiner i såvel hvede som byg har ligheder med kendte fødevareallergener og kan derfor være målproteiner i søgningen af nye allergener i f.eks. hvedemel.

CONTENTS

Chapter 1 — Introduction of *N*-glycosylation in plants and Glycoproteome analysis

ABBREVIATIONS

AIA	*Artocarpus integrifolia*
CID	Collision-induced dissociation
ConA	Concanavalin A
DBA	*Dolichos biflorus*
DSA	*Datura stramonium*
ECA	*Erythrina cristagalli*
ECD	elctron-capture dissociation
Endo H	Endoglycosidase H
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ESI	Electrospray ionization
ETD	Electron-transfer dissociation
FT	Fourier transform ion cyclotron
GA	Gibberellic acid
GSL I	*Griffonia simplicifolia* Lectin I
GSL II	*Griffonia simplicifolia* Lectin II
HILIC	Hydrophilic interaction chromatography
HRP	Horseradish peroxidase
IGOT	Isotope-coded glycosylation site-specific tagging
LCA	*Lens culinaris* Agglutinin
LC-MS	Liquid Chromatography Mass Spectrometry
LEA	*Lycopersicon esculentum*
MALDI	Matrix assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
OST	Oligosaccharyltransferase
PHA-E	*Phaseolus valgaris* Erythreoglubulin
PHA-L	*Phaseolus valgaris* Leucoagglutinin
PNA	Peanut Agglutinin Lectin
PNGase A	Peptide *N*-glycosidase A
PSA	*Pisum sativum* Agglutinin
RCA-I	*Ricinus communis* Agglutinin I
SBA	Soybean Agglutinin
SCX	Strong cation exchange
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SJA	*Sophora japonica*
STA	*Solanum tuberosum*
SWGA	Succinylated Wheat Germ Agglutinin
TFA	Trifluoroacetic acid
TOF	Time-of-flight

TOF-TOF	Tandem Time-of-flight
UEA-I	Ulex europaeus Agglutinin I
VVA	*Vicia villosa*
WGA	Wheat Germ Agglutinin
ZIC-HILIC	Zwitter ionic–hydrophilic interaction liquid chromatography

1.1 *N*-Glycosylation in plants

Glycosylation is a major post-translational protein modification and the glycan moiety on a protein is known to affect protein folding, stability, activity, interactions and functional properties (Spiro et al., 2002). The majority of proteins synthesized in the rough endoplasmic recticulum (ER) undergo glycosylation. Glycosylations can be either *N*- or *O*-linked. *N*-linked glycosylation is the attachment of a sugar molecule *via* an amide bond to the nitrogen atom in the side chain of an asparagine (N) residue in a protein. The *N*-linked glycosylation process occurs widely in eukaryotes but very rarely in bacteria. *N*-glycosylation is the most studied glycosylation event (Fitchette et al., 2007). Unlike *N*-linked glycosylation, *O*-linked glycosylation is processed by the attachment of a sugar molecule to an oxygen atom in the side chain of serine (S) and threonine (T) amino acid residues. In this thesis, the main focus is on *N*-glycosylation.

1.1.1 *N*-glycosylation pathway

The *N*-glycosylation pathway is shown in Figure 1. The process starts in the endoplasmic reticulum (ER) by transfer of an oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) from a dolichol lipid anchor at the ER membrane onto a nascent polypeptide at a specific asparagine (N) residue in an N-X-S/T/C consensus sequence (Pless and Lennarz, 1977). X can be any amino acid except proline (P). This glycosylation reaction is catalyzed by an oligosaccharyltransferase (OST) complex (Yan and Lennarz, 1999). Seven and nine protein components of the OST complex have been found in mammals and yeast, respectively (Knauer and Lehle, 1999). In plants, sequence homology searches against the *Arabidopsis* genome suggested that seven genes encoding a putative OST complex are present (Gallois et al., 1997) but the function of only three of these components have been characterized (Gallois et al., 1997, Koiwa et al., 2003, Lerouxel et al., 2005). The signal peptide at the *N*-terminus of a secreted protein is required for translocation of newly synthesized proteins into ER through the ER-translocon (Sec61). The majority of proteins secreted into ER are indeed modified by *N*-linked glycans (Helenius and Aebi, 2004), but glycosylation of potential *N*-glycosylation sites is however not mandatory (An et al., 2009). It is believed that several factors are additionally required for *N*-glycosylation such as flanking peptide sequence, the rate of protein folding and the capacity of the OST glycosylation machinery (Apweiler et al., 1999, Petrescu et al., 2004).

During transportation along the secretory pathway, the *N*-linked oligosaccharide undergoes several maturation steps starting by removal of glucose (Glc) and mannose (Man) residues. Trimming of the distal α-1,2 glucose and two adjacent α-1,3 linked glucose residues is catalyzed by glucosidase I and glucosidase II respectively (Grinna and Robbins, 1979, Hubbard and Ivatt, 1981). In plants, GCS1 encoding an Arabidopsis α-glucosidase I homologue has been cloned and characterized. The encoded protein has homology to a protein which trims the distal α-1,2 linked glucose in animal and yeast (Boisson et al., 2001). Plant glucosidase II was first identified in potato (Taylor et al., 2000) and the gene encoding this enzyme was found in a study of temperature sensitive mutant (rsw3) (Burn et al., 2002). After the glucose trimming step, mannosidase I is responsible for trimming one to four α-1,2-mannose residues in ER (Szumilo et al., 1986a, b). The addition of an *N*-acetylglucosamine moiety to the 1,3-mannose by *N*-acetylglucosamine transferase I (GnT1) is required for synthesis of complex-type *N*-glycans. These modification steps in ER and the Golgi apparatus are conserved in higher eukaryotes (Faye et al., 2005).

The pathways of late glycan maturations in the Golgi apparatus resulting in the generatation of complex-type *N*-glycans differ in plants and mammals (Figure 1). In mammals, core α-1,6-linked fucose and terminal sialic acids (NeuAc) are attached, whereas core bisecting β-1,2-xylose and core α-1,3-fucose are found in plants (Song et al., 2011).

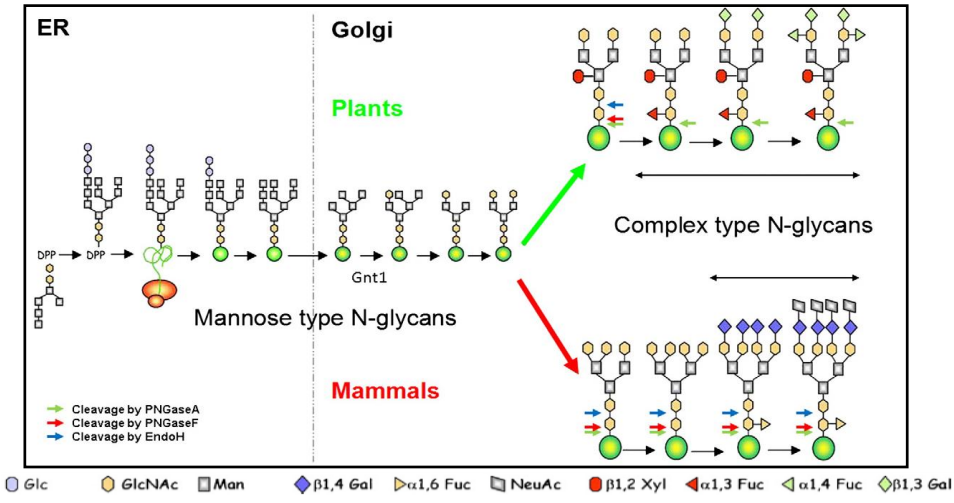


Figure 1 *N*-glycosylation pathway in plants and mammals (Song et al., 2011).

1.1.2 Roles of *N*-glycosylation in plants

In mammals, *N*-glycosylations have been proven to be involved in several biological functions such as receptor binding, cell signaling, protein folding, subcellular distribution and localization, protein stability, endocytosis, immune recognition, inflammation and pathogenicity (reviewed by Song et al., 2011). In the ER, *N*-glycan also serves as tag for marking the quality of the protein folding state (Helenius and Aebi, 2004). Most misfolded proteins are still retained in ER and degraded by ER-associated degradation (ERAD) (Hampton, 2002).

Formation of the dolichol oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) is an essential step in *N*-glycosylation pathway. Tunicamycin, an inhibitor of the enzyme GlcNAc phosphotransferase (GPT) which catalyzes the biosynthesis of the dolichol oligosaccharide precursor, has been used to block *N*-glycosylation. Studies with this antibiotic indicate that some proteins require *N*-linked oligosaccharides in order to fold properly in the ER.

Currently available information on glycosylation of plant proteins is limited, as compared to mammals. However, it has been reported that glycosylation influence enzyme activity (Kimura et al., 1999, Lige et al., 2001), thermostability and folding (Lige et al., 2001), oligomerization (Kimura et al., 1999), host-pathogen interaction (Lim et al., 2009) or subcellular localization and secretion (Ceriotti et al., 1998).

Recently, mutants of mannosidase encoding genes responsible for *N*-glycan trimming activity showed short root, with radially swollen cortical cells and alteration of cell wall development in *Arabidopsis thaliana* (Liebminger et al., 2010). This suggested that several glycoproteins are involved in cell wall formation (Liebminger et al., 2010).

A special *N*-glycan structure known as ‘Lewis a antigen’ has been also found in some plant glycoproteins (Fitchette et al., 1997, Melo et al., 1997). This structure is located at the cell surface and is involved in cell-cell recognition and adhesion in mammals (Lerouge et al., 1998). In plants, the Lewis glycan structure is also found at cell surfaces, but its functional importance has not yet been demonstrated.

Lack of complex glycans is linked to severe phenotypes in mammals for example type II congenital disorders (Schollen et al., 2005). However, little or no effect has been observed in plant mutants lacking complex glycans. The glycosylation mutant cgl that is lacking a

functional *N*-acetylglucosaminyltransferase I (GnTI) gene has been produced in *A. thaliana*. This enzyme is responsible for adding an *N*-acetylglucosamine moiety to the 1,3-mannose, which is the critical step for complex glycan formation (Von Schaewen et al., 1993, Strasser et al., 2005). However, no visible effect on the growth phenotype of *A. thaliana* was observed except a higher sensitivity to salt stress (Kang et al., 2008).

Core bisecting β-1,2-xylose and core α-1,3-fucose which are found in plants but not in mammals (Figure 1) are also known to be an important human IgE binding carbohydrate determinant of plant allergens (Bardor et al., 2003) and this also causes a problem when using plant-made proteins for therapy.

1.2 Mass spectrometry and Glycoproteome Analysis

1.2.1 Mass spectrometry

Mass spectrometry is an analytical technique used for measuring mass-to-charge (m/z) ratios of molecules and is widely applied on proteins, peptides, carbohydrates, DNA, drugs, and many other biologically relevant molecules. It has also become one of the most comprehensive tools in proteomics for characterization, identification, sequencing and structure elucidation of proteins in complex mixtures. In mass spectrometry, ions are separated according to their mass-to-charge ratios. A mass spectrometer consists of three major units, an ion source, an analyser and a detector (Aebersold and Mann, 2003). Ion sources based on matrix assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) techniques are most commonly used for analysis of proteins or peptides. In MALDI analytes in the solid state are embedded in a crystalline matrix (Figure 2a). Upon laser irradiation the matrix molecules and analytes are sublimated and mainly singly charged ions are produced. Unlike MALDI, the ESI ion source produces mainly multiply charged ions from a solution in a metal capillary. The ionization is driven by a very high voltage (2-6 kV) which is applied to the tip of the metal capillary and the sample coming out from the tip is dispersed into an aerosol of multi-charged electrospray (ES) droplets (Figure 2b) (Banerjee and Mazumdar, 2012, Steen and Mann, 2004). MALDI-MS is generally used to analyse simple peptide mixtures, whereas coupling of liquid-chromatography with an ESI-MS system (LC-MS) is preferred for the analysis of complex

samples due to the chromatographic separation reduces the complexity prior to MS analysis.

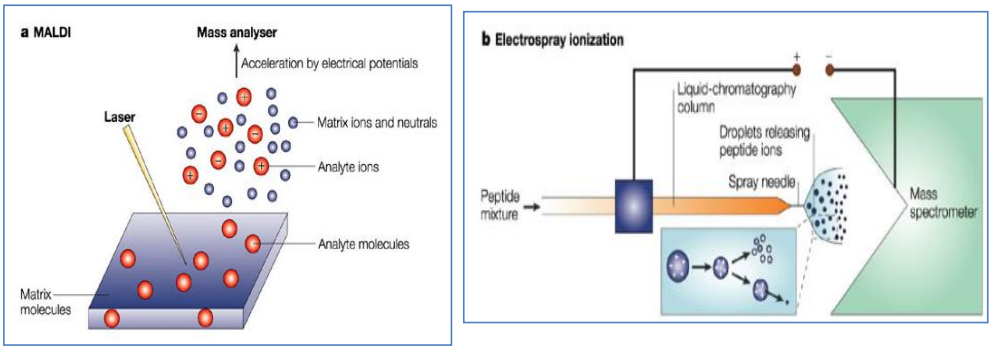


Figure 2. Ionization techniques for mass spectrometry a) In matrix assisted laser desorption ionization (MALDI), the analyte is mixed with a large excess of ultraviolet-absorbing matrix, the excess matrix molecules sublime and transfer the embedded non-volatile analyte molecules into the gas phase and singly charged analyte ions are formed. b) In electrospray ionization (ESI) the ionization is driven by a very high voltage which is applied to the tip of the metal capillary. The peptides emerging from the capillary is thereby electrostatically dispersed and generates highly charged droplets. Very small droplets are produced due to repetitive droplet fission (Steen and Mann, 2004).

The function of the mass analyser is to separate or resolve the ions emerging from the ionization source according to mass-to-charge (m/z) ratio. There are several types of mass analysers currently available for example time-of-flight (TOF), ion-trap, orbitraps, quadrupoles and Fourier transform ion cyclotron (FT) analyser (Figure 3). Tandem mass spectrometers (MS/MS) are instruments that have more than one mass analyser. Ion precursors selected from the first mass analyser are typically fragmented and subsequently resolved in the second mass analyser. Examples of tandem mass spectrometers include tandem time-of-flight (TOF-TOF), triple quadrupole (QqQ) and quadrupole-time-of-flight (Q-TOF) (Figure 3b, c and d respectively).

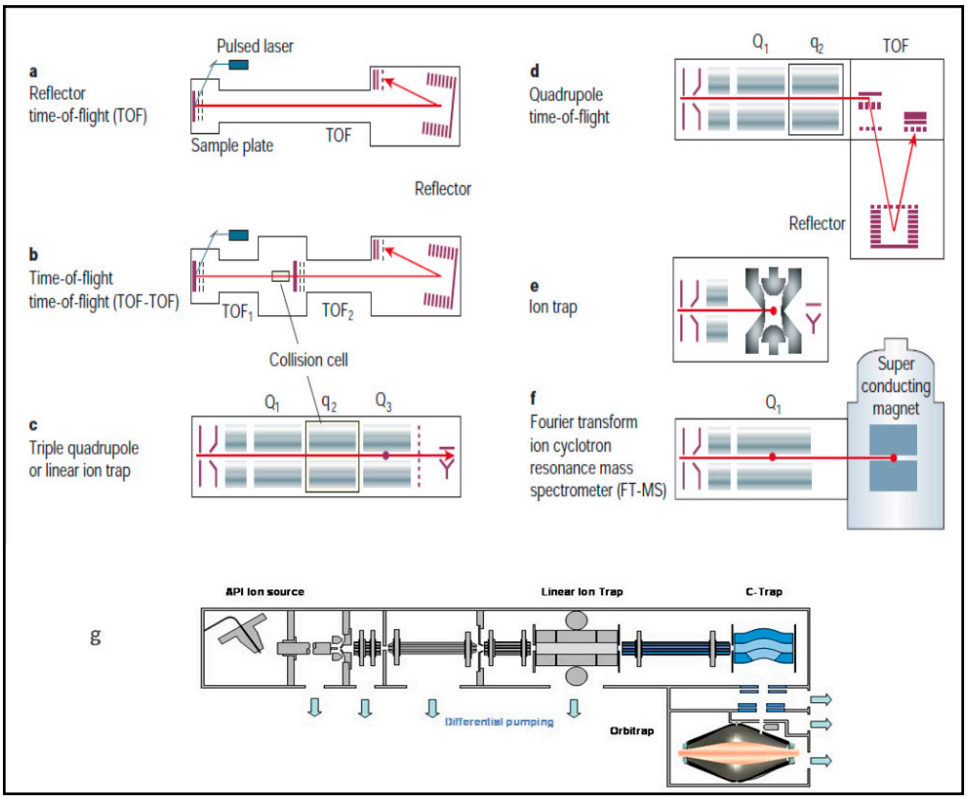


Figure 3. Different mass analysers in mass spectrometry a) In reflector time-of-flight (TOF) instruments, ions accelerated to high kinetic energy are separated along a flight tube as a result of their different velocities. b) In TOF-TOF instruments, ions of a defined mass-to-charge (m/z) ratio separated in the first TOF are fragmented and re-accelerated into the second TOF. c) In quadrupole mass spectrometers ions are selected by time-varying electric fields between four rods, which permit a stable trajectory only for ions of a particular desired *m*/*z*. Ions of a particular m/z are selected in a first section (Q_1), fragmented in a collision cell (q_2), and the fragments separated in Q_3. d) The quadrupole TOF instrument combines the front part of a triple quadruple instrument with a reflector TOF section for measuring the mass of the fragment ions. e) In ion traps, ions of a particular *m*/*z* captured in an electric field are fragmented and ejected in a *m*/*z*-dependend manner to generate the tandem mass spectrum. f) In the FT-MS instrument ions are trapped with the help of strong magnetic fields. If ion velocity is low and the field is intense, the radius of the trajectory becomes small and ion can thus be trapped. The figure shows the combination of FT-MS with the linear ion trap for efficient isolation, fragmentation and

fragment detection in the FT-MS section. g) In orbitrap, an electric field is set between an outer barrel-like electrode and an inner axial electrode. The ions end up moving in ring-like orbits around the axial electrode. Once their attraction to the inner electrode is exactly balanced by the centrifugal forces, the ions are trapped in their orbits. The signal of these oscillations can be detected on the orbitrap outer electrodes and is amplified, transformed into a frequency spectrum by fast Fourier transformation and finally converted into a mass spectrum, where the oscillation frequency for each ion is used to calculate its m/z ratio. The figure shows the combination of orbitrap with the linear ion trap (figure adapted from www.thermo.com and Aebersold and Mann, 2003).

There are several methods used to fragment ions in mass spectrometer, for example collision-induced dissociation (CID), electron-transfer dissociation (ETD) and electron-capture dissociation (ECD). In collision-induced dissociation (CID), the ion precursors are usually accelerated by electrical potentials to high kinetic energy and then allowed to collide with neutral molecules (often helium, nitrogen or argon). The collision of peptides results in bond breakage and the fragmentation of the peptide ion precursors into smaller fragments (Roepstorff and Fohlman, 1984). Electron-transfer dissociation (ETD) induces fragmentation of cations (e.g. peptides or proteins) by transfering electrons to them. The ETD fragmentation of a peptide occurs randomly along the peptide backbone generating mainly c- and z- type ions (Syka et al., 2004). In general, ETD is particularly effective for highly charged peptides. Being able to cope with highly charged peptides, ETD promotes the use of alternative proteases such as Lys-C, Lys-N, or Arg-C. During ESI, these peptides result in parent ions having three to six charges which make them ideal candidates for ETD-mediated fragmentation (Wiesner et al., 2008). Electron-capture dissociation (ECD) fragment gas phase ions by direct introduction of low energy electrons to trapped gas phase ions. ECD produces different types of fragment ions primarily c-, z- and b type ions. In contrast to CID, post-translational protein modifications such as phosphorylations and glycosylations are typically retained on peptide fragments generated by ETD and ECD.

After the mass analyser(s) the ions reach the detector where the signal is amplified and reported in the form of mass spectra where the m/z ratios of the ions are plotted against their signal intensities.

In MALDI, fragmentation of ions occuring in the TOF mass analyser is referred to as post-source decay (PSD) fragmentation. It corresponds to the decomposition of metastable ions, which are stable enough to leave the source but contain enough excess energy to allow their fragmentation before they reach the detector. Since the PSD fragment ions of a particular precursor ion are all given the same kinetic energy in the MALDI source, they will reach the detector at the same time-point in a linear MALDI-TOF instrument. If the instruments is equipped with a reflectron the fragments will however be separated in the TOF according to their m/z ratio. To improve separation of PSD fragments, a second ion source cell is inserted in the ion flight path, which selects a parent ion and its fragments formed by PSD in the first TOF region. The parent and fragment ions are then re-accelerated, and therefore travel with different velocities according to their masses, and are focused on the detector after passing through the reflectron. PSD preferentially cleaves the most labile bonds (Harvey et al., 2003, Kaufmann et al., 1993, Kaufmann et al., 1994) present in the peptides for example those in phosphorylated and glycosylated peptides.

1.2.2 Glycoproteome analysis

A wealth of information is available about protein glycosylation in mammals. However, little information is available on plant protein glycosylation. The study of *N*-glycoproteomes in plant should get attention to provide more insight into various processes such as protein folding, functional activities, stability, signaling and interaction (Spiro, 2002).

Glycoproteome analysis involves the identification of glycoproteins, assignment of glycosylation sites, and determination of glycan structures and their site occupancy (Fitchette et al., 2007). Identification of glycoproteins and glycosylation site assignments will be the focus of the present chapter. *N*-glycosylation involves the attachment of *N*-glycan at a consensensus sequence NXS/T. A number of these potential glycosylation sites can be present in a protein however they may or may not be glycosylated (An et al., 2009). The verification of the potential glycosylation sites provides a basic knowledge for further detailed characterization for example by mutational analysis (Stround et al., 2001). It may also indicate the orientation of a transmembrane protein because glycan part is commonly found at the extracellular domains (Zielinska et al., 2010). In combination with existing protein structure information, the glycosylation sites may give a deeper insight into the detail structure of a protein (Pasing et al., 2012).

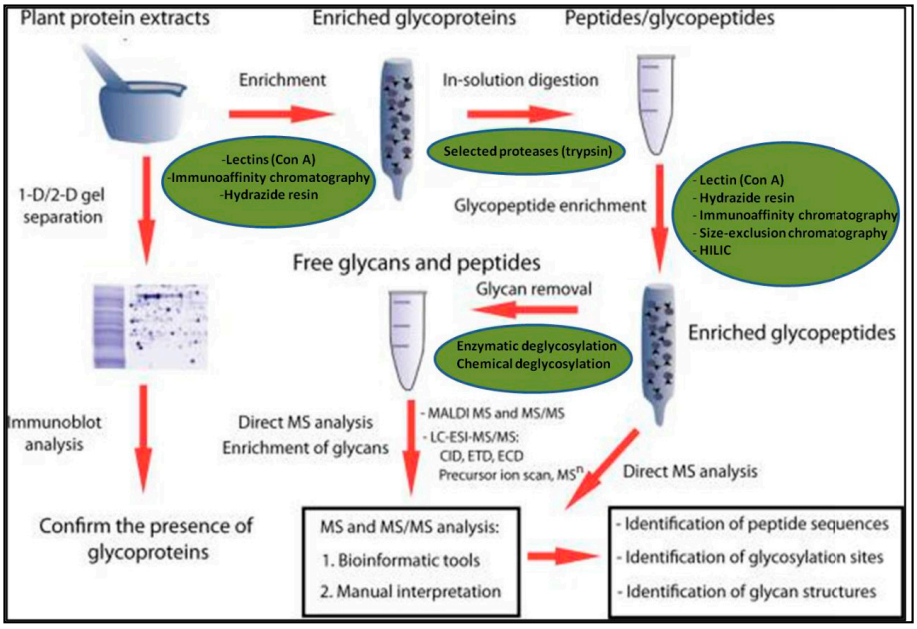


Figure 4. Workflow of the strategic glycoproteomic analysis for systematic characterization of glycoproteins by mass spectrometry. The general workflow consists of three major steps: enrichment of glycopeptides/glycoproteins, glycan removal and mass spectrometric analysis to obtain information on the identified glycoproteins, inclusive assignment of glycosylation sites, and determination of glycan structures and their site occupancy (Ruiz-May et al., 2012).

Mass spectrometry has a key role as an analytical tool for proteome analysis including protein glycosylation. However, several obstacles exist in glycosylation analysis. Glycopeptide signals are suppressed by the presence of highly abundant non-glycosylated peptides. In addition, the signal intensity of individual glycopeptides is relatively low when compared to the signal intensity of non-glycosylated peptides mainly due to large heterogeneity of the glycan structure (Hagglund et al., 2004). Glycoproteins and glycopeptides are also generally hydrophilic and are therefore often not retained in reversed phase chromatography which is commonly used for purification of peptides prior to mass spectrometry. Furthermore, only few fragment ions of the backbone of glycopeptides are usually obtained using traditional CID tandem mass spectrometry

(MS/MS) (Hogan et al., 2005; Scott et al., 2010) yielding little or no information of the peptide and glycosylation site identification due to the labile nature of the glycosidic bond.

To overcome these obstacles, several strategies have been developed and used in combination e.g. glycopeptide enrichment technologies, advanced MS instrumentions and deglycosylation of glycoprotein/glycopeptides to remove the complexity of the glycan moeties. The strategy in glycoproteomic analysis workflow is shown in Figure 4.

1.2.3 The enrichment of glycoproteins/glycopeptides

Lectin affinity

Lectins are sugar-binding proteins that interact with glycan structures attached to glycoproteins. There are several types of lectins commercial available with specificities for different types of glycans (Table 1). The binding of lectins is thought to involve mainly hydrogen bonds with some additional contribution by van der Waals and hydrophobic interactions (Sharon, 1993). Lectin affinity methods can be used for either glycoprotein or glycopeptide isolation from complex samples. Concanavalin A (Con A) is the most commonly used (Bunkenborg et al., 2004, Fan et al., 2004, Minic et al., 2007, Zhang et al., 2010, Catala et al., 2011) because of its broad specificity in recognition of oligomannosyl motifs in *N*-linked glycans (Ohyama et al., 1985). Other lectins have narrower specificity (Table 1). Because the individual lectin binds to specific types of glycan, only a subset of glycopeptides can be recognized and isolated using lectins. Use of various lectins for glycopeptides enrichment are performed sequentially (Yamamoto et al., 1995, 1998), in parallel (Yang et al., 2006; Lee et al., 2009; Zielinska et al., 2010), and as mixtures (Li et al., 2004) to overcome this limitation and hence increase the coverage of identified glycoproteins/glycopeptides.

In the first survey of a plant cell wall *N*-glycoproteome, Minic et al. used affinity chromatography on concanavalin A-Sepharose followed by 2DE-separation and identified 102 glycoproteins using nanoHPLC MS/MS and MALDI-TOF MS (Minic et al., 2007). However, glycosylation site assignments were not reported in this publication. In a study by Catala et al. Con A was applied for isolation of glycoproteins in ripe tomato fruit (Catala et al., 2011). After trypsin digestion the glycopeptides were captured with SCX. Zhang et al. recently used various lectins and boronic acid for glycoprotein enrichment

from plant cell walls followed by SDS-PAGE (Zhang et al., 2010). Altogether, 127 glycoproteins were identified.

Table 1. Available lectins and their binding specificities (Afrough et al., 2007).

Lectins	Abbreviation	Sugar Specificity
Peanut Agglutinin	PNA	Galactose
Jacalin	AIA	Galactose/GalNAc
Erythrina cristagalli	ECA	Galactose/GlcNAc
Sophora japonica	SJA	GalNAc/Galactose
Ricinus communis Agglutinin I	RCA-I	GalNAc/Galactose
Griffonia simplicifolia Lectin I	GSL I	GalNAc/Galactose
Soybean Agglutinin	SBA	GalNAc
Dolichos biflorus	DBA	GalNAc
Vicia villosa Lectin	VVA	GalNAc
Griffonia simplicifolia Lectin II	GSL II	GlcNAc
Wheat Germ Agglutinin	WGA	GlcNAc/NANA
Succinylated Wheat Germ Agglutinin	SWGA	GlcNAc
Lycopersicon esculentum	LEA	GlcNAc
Solanum tuberosum	STA	GlcNAc/Sialic acid
Lens culinaris Agglutinin	LCA	Glucose/Mannose
Concanavalin A	ConA	Glucose/Mannose
Pisum sativum Agglutinin	PSA	Glucose/Mannose
Phaseolus vulgaris Leucoagglutinin	PHA-L	Complex Sugar
Phaseolus vulgaris Erythreoglubulin	PHA-E	Complex Sugar
Datura stramonium	DSA	LacNAc
Ulex europaeus Agglutinin I	UEA-I	Fucose

The sugar specificity of the lectins and their abbreviations have been extracted from References 3 and 4. LacNAc is N-acetyl lactosamine, and NANA is N-acetylneuraminic acid.

Hydrazide chemistry

Solid phase extraction of glycoproteins based on hydrazide chemistry was first developed by Zhang et al., 2003. This protocol comprises the following key steps (Figure 5a).

1) Periodic acid oxidation of *cis*-diol groups to aldehydes (Figure 6)
2) Coupling of the aldehydes and hydrazide groups presented in the solid support to capture glycoproteins. Covalent hydrazone will be formed by this reaction.
3) Proteolytic digestion (generally by trypsin) to remove and wash away those protein fragments which are not glycosylated. The glycosylated peptides still remain on the solid support.
4) Release of glycopeptides by deglycosylation with PNGase F.

Later, this protocol was modified for the enrichment of glycopeptides by Zhou et al., 2007. Glycoproteins are first digested with trypsin followed by periodic oxidation, glycopeptide capture by coupling reaction and release of glycopeptides by PNGases respectively (Figure 5b).

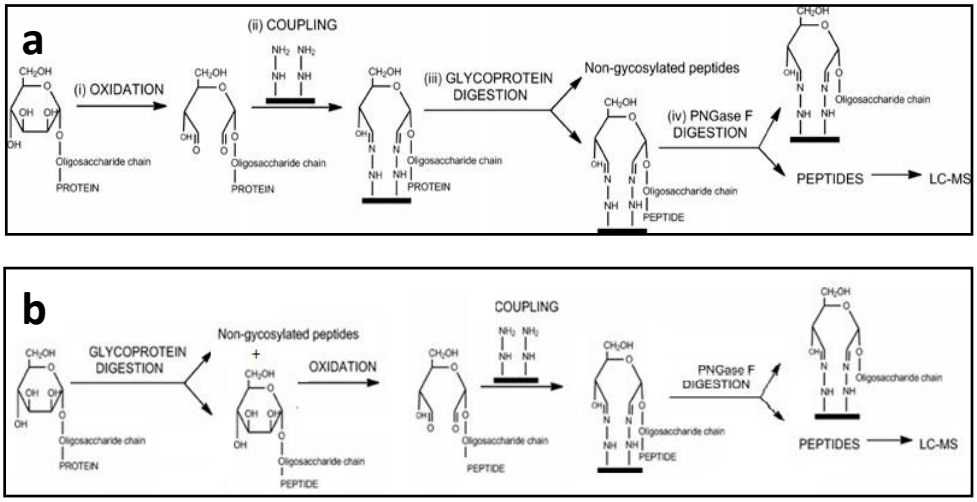


Figure 5. Hydrazide chemistry strategies for (a) glycoprotein enrichment and (b) glycopeptides enrichment (b) prior to glycosylation site analysis by LC-MS (figure modified from Ongay et al., 2012).

When comparing the enrichment of glycopeptides and glycoproteins by the hydrazide resin, the higher efficacy is shown for glycopeptides enrichment (Zhou et al., 2007, Berven et al., 2010). This may be due to more favorable access of *N*-glycopeptides to the hydrazide resin (Zhou et al., 2007, Sun et al., 2007).

For the release of glycopeptides from plant proteins, PNGase A could be applied instead of PNGase F as PNGase F is not able to cleave *N*-linked glycans where the innermost GlcNAc residue is linked to an α1-3 fucose residue. This form of *N*-glycan is the most commonly found in glycoproteins from plants and insects. The application of PNGase A for plant glycopeptides release from the hydrazide resin has been demonstrated by Palmisano et al., 2010 for *N*-glycoproteome study in Chardonay white wine.

O-glycosylated proteins can also be captured by hydrazide chemistry. However, there is no enzyme to release the glycoproteins/glycopeptides from the hydrazide resin. Non-enzymatic chemical approaches such as the use of hydroxylamine are needed to release of the glycoproteins/glycopeptides from hydrazide resin (Klement et al., 2010).

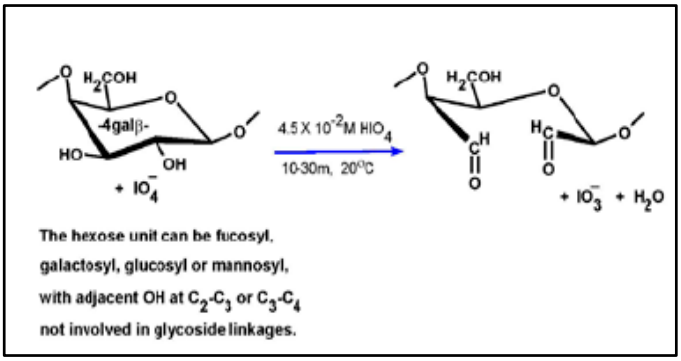


Figure 6. Periodic oxidation. Periodate ions selectively oxidixes *cis*-diol groups (a pair of adjacent hydroxyl groups in *cis* configuration), yielding two aldehyde groups.

Boronic acid chemistry

Boronic acids form cyclic esters with 1-2 and 1-3 *cis*-diol groups which are present in sugars such as mannose, galactose and glucose. These cyclic esters are stable under basic pH conditions and released at acidic pH (Figure 7). Therefore, the glycopeptide capture/release can easily be done by switching the pH from alkaline solutions in the loading step to acidic solutions in the glycopeptides elution step. Unlike lectin affinity, the recognition of a specific type of glycans is not required by boronic acid affinity. Thus, a broader range of glycoproteins/glycopeptides is obtained by this strategy. However, co-enrichment of non-enzymatically glycated proteins/peptides can be observed (Zhang et al., 2008) and the presence of high amounts of non-glycosylated peptides may also inhibit the glycopeptides-binding process (Ongay et al., 2012).

Boronic acid conjugated to matrices are commonly used for affinity chromatography. Several boronic-functionalized matrices such as monoliths (Chen et al., 2009), mesoporous

silica (Xu et al., 2009), magnetic particles or gold nanoparticles (Zhang et al., 2011, Qi et al., 2010) have been used for glycopeptide enrichment with high specificity and sensitivity.

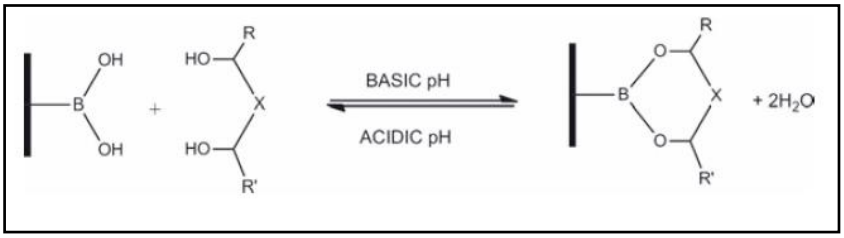


Figure 7. Boronic acid chemistry Reversible reaction of boronate and *cis*-diol-containing molecules allows glycopeptide capture/release by a pH switch (Ongay et al., 2012).

Strong cation-exchange chromatography

Strong cation exchange (SCX) chromatography has been employed for enrichment of sialylated glycopeptides (Lewandrowski et al., 2007). The presence of additional negatively charges carried on sialic acid reduces the positive net charge in comparison with non-glycosylated peptides under acidic condition (pH 2.7). Therefore, glycopeptides are eluted prior to non-glycopeptides by use of strong cation exchange resins.

Size-exclusion chromatography

N-Glycopeptides have considerably higher masses than non-glycosylated peptides because of the contribution of *N*-linked glycans. This allows the separation of glycosylated peptides from non-glycosylated peptides by size exclusion chromatography (Alvarez-Manilla et al., 2006). A small peptide that can penetrate every region of stationary phase pore system will elute late. On the other hand, a larger peptide that cannot penetrate any region of the pore system will take shorter time to elute from the column.

Hydrophilic-interaction chromatography (HILIC)

Retention in HILIC is based on the hydrophilicity of compounds. Glycopeptides are typically more hydrophilic than non-glycosylated peptides due to the presence of numerous

hydroxyl groups in the carbohydrate part, and thus ideal candidates for separation by HILIC. Different HILIC stationary phases have been used for glycopeptide separation. The mechanism of HILIC based on different types of stationary phase, efficiency, retention and selectivity is still not clearly understood (Buszewski and Noga, 2012). More information detailed information on HILIC for glycopeptide enrichment is reviewed in the next part of this chapter.

1.2.4 Gel-based glycoprotein analysis

Gel-based glycoprotein detection methods have been reviewed by Patton, 2002. The presence of glycoproteins can be visualized using 1D/2D gel-based electrophoretic separation followed by several staining techniques such as colorimetric staining (e.g. acid fushin, Alcian Blue) or fluoresescent stain (e.g. Pro-Q Emerald 488). Most staining techniques are based on the periodic acid of *cis*-diols folloed by Schiff reaction procedure. Immunoblot analysis using antibodies against glycoproteins after gel electrophoretic separation is also a possibility. Antibodies against horseradish peroxidase (HRP) can be used to detect the presence of β-1,2-xylose and/or α-1,3-fucose structures on complex glycans of plant glycoproteins (Strasser et al., 2008). In addition, the mobility shift of protein in gel electrophoresis after deglycosylation can be used to confirm the presence of glycoproteins (Fitchette et al., 2007).

1.2.5 Mass spectrometric analysis of intact glycopeptides

Traditionally, collision-induced dissociation (CID) is used to fragment peptides in tandem mass spectrometry. This type of fragmentation technique provides information about the seqeunce of the peptide since fragmentation often occurs across peptide bonds in the backbone. Unfortunately, glycosidic bonds are however very labile and the glycans are not present in the observed fragments (Hogan et al., 2005, Scott et al., 2010).

Recently, electron transfer dissociation (ETD) (Syka et al., 2004) and electron capture dissociation (ECD) (Mirgorodskaya et al., 1999) were introduced for the characterization of several types of protein post translational modifications including glycosylation (Pasing et al., 2012, Wuhrer et al., 2007). Peptide fragmentation by ETD and ECD cleaves N-Cα bonds of the peptide backbone generating c- and z-type fragment ions without loss of glycan moieties on glycosylation sites. Because glycan moieties remain intact, it

subsequently enables the identification of both the glycosylation site and the glycan moiety (Figure 8a).

However, there are some limitations of ETD/ECD based fragmentation for the characterization of glycopeptides for example the fragmentation efficiency of ETD is low. For glycopeptides, the fragmentation efficiency is approximately 20% (Mechref et al., 2012). ETD appears to be limited to an m/z range of less than about 1400 (Swaney et al., 2007) and current database search algorithms are still not readily compatible with ETD-based approaches (for review see Pasing et al., 2012).

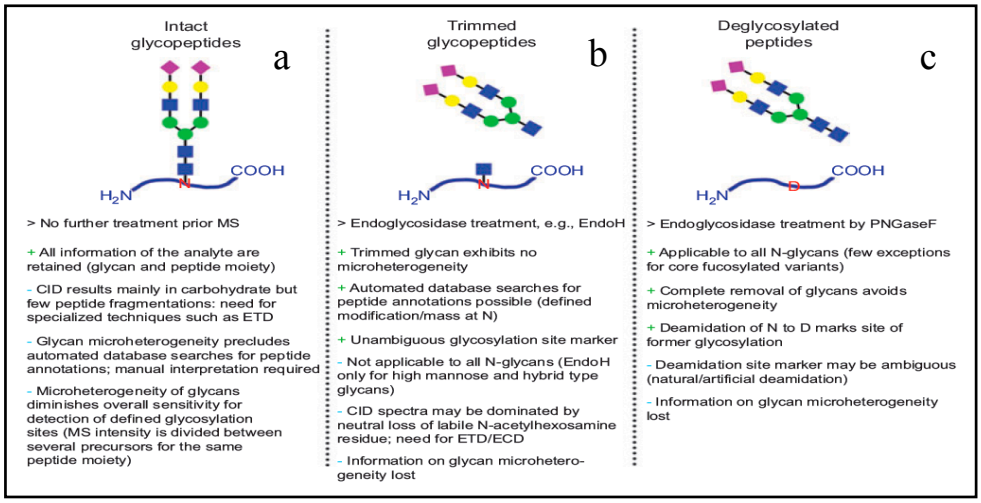


Figure 8. MS-based *N*-glycosylation site workflows using (a) ETD/ETC Tandem MS of intact glycopeptides or removal of glycan moiety by (b) Trimming with EndoH (c) Deglycosylation with PNGase F/PNGaseA (Pasing et al., 2012).

1.2.6 Removal of glycan moieties of glycopeptides by PNGase-deglycosylation and Isotope-coded glycosylation site-specific tagging (IGOT)

In order to facilitate interpretation of glycopeptide mass spectra, *N*-glycan moieties are removed prior to mass spectrometric analysis in most workflows. Removal of glycan moieties improves the efficiency of CID type fragmentation of glycopeptide backbone and

also increases the signal intensity of glycopeptides because of the reduced heterogeneity (Wuhrer et al., 2007).

PNGase F has been intensively used in glycopeptide identification workflows in mammals (Kaji et al., 2006, Fan et al., 2004). This enzyme cleaves *N*-glycans between the innermost GlcNAc and the asparagine residues (N) of high mannose, hybrid, and complex oligosaccharides from *N*-glycoproteins (Figure 8c). However, it is not able to hydrolyze *N*-linked glycans from glycoproteins in case the innermost GlcNAc residue is linked to an α-1,3-fucose residue as commonly found in plant.

PNGase A is the enzyme of choice for use in plant glycopeptide identification. The enzyme was first isolated from almonds and cleaves *N*-glycan with or without an α-1-3 fucose residue attached to the innermost GlcNAc residue. Sialylated glycans, commonly found in mammals, can also be hydrolyzed by this enzyme. However, PNGase A is not efficient on intact glycoproteins even if they have been denatured. (Plummer et al., 1981, Tretter et al., 1991) therefore proteolytic digestion is required for application of PNGase A.

Deglycosylation reactions using PNGase A and F cause deamidation of the asparagine residue which resulting in convertion from the asparagine to aspartate. This reaction results in a mass increase of 0.984 Da. Therefore the deamidation can be used for differentiation of Asn at the glycosylation site from other Asn residues leading to assignment of formerly *N*-glycosylated sites (Figure 8c). However, Asn deamidations do not occur only due to deglycosylation reactions but are also generated from spontaneous deamidation *in vitro* or natural biological processes *in vivo*. This ambiguity can lead to false-positive glycosylation site assignment (Parker et al., 2011).

In order to minimize false positive identifications, the deglycosylation reaction using PNGase A and F can be performed in the presence of $H_2{}^{18}O$. This approach is termed "isotope-coded glycosylation site-specific tagging (IGOT)" (Kaji et al., 2003). This approach facilitates identification of glycosylation sites since the ^{18}O atom is incorporated into the carboxylate group of deamidated asparagine residues during the PNGase reaction whereas residues that were deamidated before the enzymatic reaction was initiated contains ^{16}O only. By PNGase-deglycosylation, the conversion of glycosylated asparagine residues into ^{18}O-labeled aspartic acids will occur and result in 2.9883 Da mass increase due to the incorporation of ^{18}O (Kaji et al., 2003, Ueda et al., 2010, Beck et al., 2011)

(Figure 9). This is readily distinguished from the 0.984 Da mass shift observed when the deamidation reaction previously takes place in $H_2{}^{16}O$. However, this approach cannot exclude the possibility of incorporation of ^{18}O into Asn residues due to spontaneous chemical deamidation during the deglycosylation process. Many chemical factors can influence the rate of spontaneous deamidation for example alkaline pH, buffer component and temperature (Pasing et al., 2012). In addition, an asparagine residue close to glycine was often found to be deamidated. Hao et al., 2011 demonstrated that decreased pH value for deglycosylation (pH 5) could reduce spontaneous deamidations. However, the pH optimum of PNGase A, which is commonly used for plant glycopeptides is pH 4.0-6.0, whereas the pH 7.5 is commonly used for PNGase F (recommended by the manufacturer).

Furthermore, even after reversed-phase clean up and heat denaturation, the trypsin used for proteolysis can still be active and lead to incorporation of ^{18}O into the C-termini of peptides (Kaji et al., 2006). Such labelling at the C-terminus may result in a high rate of false positive identifications (Angel et al., 2007) as it is difficult to distinguish the mass increase from incoporation of ^{18}O to C-terminus and glycosylation sites. Several approaches have been proposed to overcome this problem for instance by removing trypsin or other proteases prior to deglycosylation, or by changing to a buffer prepared in $H_2{}^{16}O$ after completion of the PNGase treatment (to change ^{18}O to ^{16}O at the C-terminus), or by setting parameter allowing ^{18}O labeling at the C-terminus during database search (Angel et al., 2007).

Another choice for removal of glycan moieties in plants is by trimming off the major part of the glycan by using endoglycosidase H (Endo H) or D (Endo D). These enzymes cleave glycosidic bonds between the two GlcNAc residues in the core chitobiose unit (Pasing et al., 2012, Hägglund et al., 2004) and leaves the innermost GlcNAc remaining on the glycopeptide chain (Figure 8b). High-mannose and hybrid type *N*-glycans can be trimmed from glycoproteins with Endo H. Endo D is active on complex type glycans when used in combination with a combination of exoglycosidases (Koide and Muramatsu, 1974). The *N*-glycan bond between asparagine and the remaining innermost GlcNAc residue is more stable than O-glycosidic bonds and enables assignment of glycosylation sites (Hagglund et al., 2004, Pasing et al 2012). However this *N*-glycan bond is less stable than peptide bonds and may occationally cause problems with interpretation of CID fragmentation spectra (Wiesner et al., 2008).

Chemical deglycosylation can also be used to remove *N*- and *O*-linked glycans, however, it often generates unexpected modifications of peptide side chains (Fitchette et al., 2007, Sojar et al., 1987).

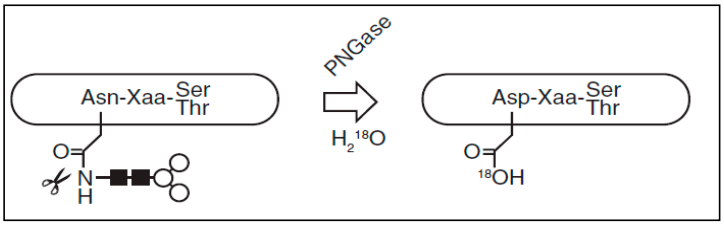


Figure 9. Isotope-coded glycosylation site-specific tagging (IGOT) strategy to discriminate PNGase-deamidated glycosylation sites from deamidation sites present prior to the onset of deglycosylation. The conversion of glycosylated asparagine residues into ^{18}O-labeled aspartic acids result in 2.9883 Da mass increase due to the incorporation of ^{18}O (Kaji et al., 2006).

1.3 Glycopeptide enrichment by Hydrophilic Interaction Liquid Chromatography (HILIC)

1.3.1 HILIC for glycopeptide enrichment

Hydrophilic Interaction Liquid Chromatography (HILIC) is a separation technique which targets hydrophilic/polar molecules by using a hydrophilic stationary phase and a hydrophobic mobile phase. Its mechanism is considered to be the opposite of reversed phase chromatography (Alpert et al., 1990, Buszewski and Noga., 2012). Various polar molecules such as nucleic acids, amino acids, and carbohydrates including glycopeptides usually have little or no retention in reversed phase chromatography, but strong retention on HILIC column (Guo and Gaiki, 2005, Boersemam and Mohammed, 2008, Alpert et al., 1990). The HILIC technique was first described by Linden et al. 1975. HILIC can be characterized as a subtype of normal-phase chromatography (NPC). The acronym HILIC was used to distinguish it from NPC; non-aqueous solvents were typically used in NPC, while water was added in eluents for the separation by HILIC (Alpert et al., 1990). The addition of water improves solubility of polar analytes.

HILIC has been extensively used in various applications. In recent years, HILIC has been applied in various proteomic studies for example in multidimensional peptide separation and analysis of post-translational protein modifications in the targeted analysis of phosphorylation, glycosylation, N-terminal acetylation and histone modifications (Boersema and Mohammed, 2008). In this thesis I will focus on the use of HILIC in glycopeptide analysis.

Retention of glycopeptides on HILIC columns is based on hydrophilic interactions with hydroxyl groups in the attached glycans (Figure 10). In case of sialylated glycans, negatively charges also confer electrostatic repulsion interaction with some HILIC materials. The HILIC retention of glycans in general may be due to hydrogen bonding, ionic interactions and dipole-dipole interactions (Wuhrer et al., 2009). For the HILIC separation of glycopeptides, not only attached glycans but also the hydrophilicity of glycopeptides depends on their core amino acid components. While the enrichment of glycopeptides by lectin affinity chromatography or hydrazide chemistry capture provides a subset of glycopeptides due to their inherent specificities, HILIC can be utilized for a broad range of glycopeptides.

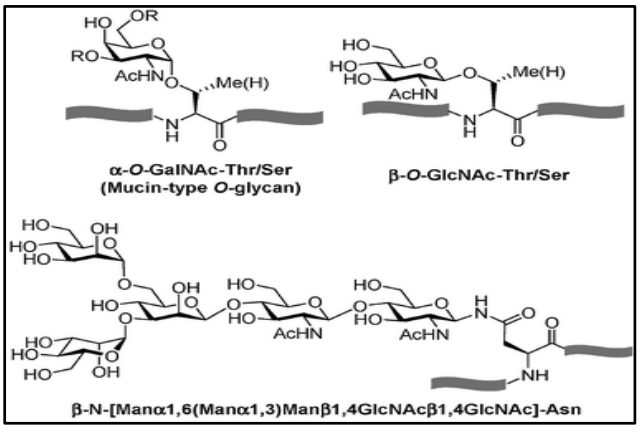


Figure 10. Abundance of hydroxyl groups in carbohydrate moieties of glycopeptides makes them highly hydrophilic and results in retention on HILIC columns (Park et al., 2008).

1.3.2 The principle of HILIC retention

Although the HILIC technique has been well established since 1975, the exact mechanism of HILIC retention is still debated. The mechanism is believed to be based on partitioning of analytes between the bulk organic eluent and the water enriched layer that is absorbed on the surface of the polar stationary phase (Alpert et al., 1990). This is in contrast to the retention in normal phase (NP) where analytes are retained by absorbtion onto the stationary phase. In HILIC, the more hydrophilic the analyte is, the more it partitions in the part of immobilized water layer that surrounds the stationary phase, and thus highly hydrophilic analytes can be immobilized on the column while less hydrophilic compounds can be separated and released to the flow-though by an hydrophobic (organic) mobile phase (Buszewski and Noga, 2012). An increase in percentage of organic solvent (e.g. acetonitrile) also causes a stronger interaction of water with the stationary phase. The HILIC partitioning process is illustrated in Figure 11. In addition, electrostatic interaction is considered as a secondary interaction for the retention of charged analytes separated by some types of stationary phases (Hemstrom and Irgum, 2006).

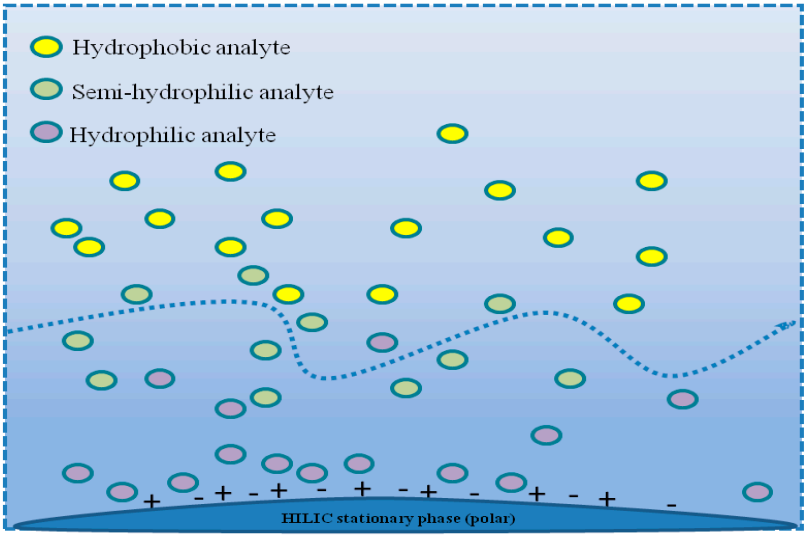


Figure 11. Principle of HILIC retention. HILIC retains a water-enriched layer on the stationary surface. This water layer facilitates partitioning of polar analytes from the organic mobile phase to the hydrophilic stationary phase resulting in increased retention of polar compounds.

In summery, the retention on the HILIC stationary phase depends on various types of attractive interactions between the analyte and the stationary phase, the analyte and the mobile phase, and the stationary phase and mobile phase (Buszewski and Noga, 2012).

1.3.2.1 Stationary phase

HILIC stationary phases are made from polar chromatographic materials. Many types of HILIC stationary phase materials are currently available. Stationary phases can be classified into three different catagories based on electrostatic interactions contributed from HILIC materials, neutral, charged and zwitterionic stationary phases. So far, several types of polar materials have been utilized for enrichment of glycopeptides. Some selected materials are listed in Table 2 and structures of some selected HILIC materials are shown in Figure 12.

Table 2. Selected HILIC materials reported for glycopeptide isolation.

HILIC material	References
ZIC-HILIC	Hagglund et al., 2004
Click maltose	Yu et al., 2009
Sepharose and microcrystalline cellulose	Wada et al., 2004
Cotton wool microcolumns	Selman et al., 2011
Click chitooligosaccharide	Huang et al., 2009
Click novel glycosyl amino acid	Huang et al., 2011
TSKgel Amide-80	Parker et al., 2011
HILIC XBridge column	Pompach et al., 2012
Amine-functionalized magnetic nanoparticles	Kuo et al., 2012
Click OEG-CD matrix	Zhao et al., 2011
Zirconia layer coated mesoporous silica microspheres	Wan et al., 2011
Magnetic bead-based zwitterionic HILIC	Yeh et al., 2012

Several carbohydrate-based stationary phases are used for glycopeptide enrichment for instance sepharose, microcrystalline cellulose (Wada et al., 2004), click maltose (Yu et al., 2009) and click chitooligosaccharide (Huang et al., 2011). Recently, cotton wool has also been successfully applied for the selective enrichment of glycopeptides from tryptic digest of human IgG and fetuin (Selman et al., 2011). The cotton wool micro-columns are cheap and very easy to prepare. It also gives fast flow velocity during the separation. In this thesis, the first report applying cotton wool for glycopeptides isolated from a complex biological sample is described. These carbohydrate-based materials are considered as neutral stationary phases. The presence of hydroxyl groups on carbohydrates forms the hydrophilic properties of the stationary phases (Figure 12). Theoretically, HILIC retention of the carbohydrate-based solid phase is solely based on hydrogen bonding (Selman et al.,

2011) by partitioning of hydrophilic glycopeptides between a water layer and the hydrophobic eluent.

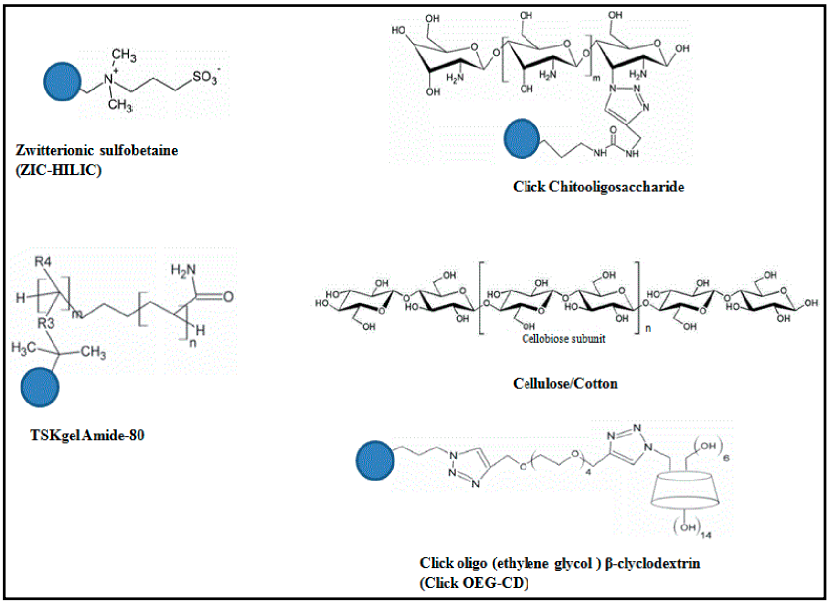


Figure 12. Structures of some selected HILIC materials.

Stationary phases with zwitterionic sulfobetaine functional groups so-called zwitterion chromatography–hydrophilic interaction chromatography (ZIC–HILIC) materials are also applied for glycopeptide enrichment. ZIC-HILIC materials are zwitterionic and thus contain both negative and positive charges (Figure 12). ZIC-HILIC resin is one of the most commonly used stationary phases for HILIC solid phase extraction (SPE) of glycopeptides. The retention on ZIC-HILIC is considered to be based on polar and weak electrostatic interactions. By comparison of different common HILIC materials (bare silica, microcrystalline cellulose, TSKgel Amide-80, and sulfobetaine-(ZIC-HILIC)) as demonstrated by Wohlgemuth J et al., 2009, ZIC-HILIC showed the highest selectivity and high signal intensity compared to noise for glycopeptides. For making a SPE microcolumn, the limitation is high back pressure during separation due to its very fine particle size (10

μm). Therefore only small amounts ZIC-HILIC resin can be loaded in microcolumns resulting in lower capacity for peptides binding. In the present thesis, this limitation is overcome by using ZIC-HILIC in combination with cotton for larger scale enrichment and fast flow separation.

Many recently developed HILIC SPEs are synthesized by click chemistry by linking a functional group, often a saccharide, to silica or nanoparticles for example maltose, chitooligosaccharide, novel glycosyl amino acid and cyclodextrin (CD) bonded silica. Underivatized silica (bare silica) itself is considered to be hydrophilic due to silanol groups on the surface of silica being hydrophilic. However, little or no retention is obtained when glycopeptide isolation is done by bare silica (Wohlgemuth et al., 2009). The electrostatic interaction on bare silica is dependent on the pH of the mobile phase. Generally, it is assumed that all silanol groups are fully protonated at pH 4.0 and that the silanol groups start to ionize above pH 4.0 (Guo and Gaiki, 2005).

1.3.2.2 Mobile phase and utilizing ion-pairing for glycopeptide enrichment by HILIC

Typically, the mobile phase for HILIC separation contains water-miscible organic solvents and water. When the concentration of organic solvent increases polar analytes are more strongly retained (Buszewski and Noga, 2012).

The most popular mobile phase composition is acetonitrile and water. Other water-miscible organic solvents can be utilized, for instance, tetrahydrofuran, isopropanol, and methanol. The choice of mobile phase composition has great effects on retention and selectivity on some samples. The strengths of some solvents are summerized as follows:

Acetone < isopropanol ~ propanol < acetonitrile < ethanol < dioxane < DMF ~ methanol < water (Buszewski and Noga, 2012)

Due to the pH-dependency of the charge of some stationary phases and analytes, the pH of the HILIC mobile phase is commonly adjusted by addition of formic acid and/or ammonium acetate (Buszewski and Noga, 2012).

In the separation of glycopeptides by HILIC, hydrophilic non-glycosylated peptides can also be co-enriched. This co-enrichment arises from an overlap of the hydrophilicity of glycopeptides and non-glycosylated peptides. Ding et al., 2009 demonstrated that the addition of ion pairing agent such as trifluoroacetic acid (TFA) to the mobile phase

decreases the hydrophilic overlap and improves the separation of glycopeptides from non-glycosylated peptides by HILIC ESI-MS. Recently, Mysling et al., 2010 also demonstrated the use of ion-pairing agent (TFA) for decrease in co-enrichment of non-glycosylated peptides and subsequent improvement of glycopeptides detection by ZIC-HILIC SPE packing in a microcolumn.

1.4 Cereal grain of wheat and barley

Cereals are the edible seeds of plants in the Gramineae family (grass family). Wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.) belong to the tribe Triticeae within the Pooideae subfamily of grasses. The taxonomy of wheat and barley is shown below (NCBI taxonomy browser, www.ncbi.nlm.nih.gov/Taxonomy/Browser/)

Kingdom:	Viridiplantae (Plant)
Phylum:	Streptophyta (Green plant)
Clade:	Liliopsida (Monocots)
Order:	Poales
Family:	Poaceae
Subfamily:	Poaceae
Tribe:	Triticeae

Wheat and barley grains are widely used for human consumption and animal feed. Food products from wheat include for example bread, pasta, biscuits and cake. Barley is used mainly in the brewing industry which is of great economic importance. Some kind of beer is also produced from wheat grains. Wheat and barley grains predominantly contain 60% to 70% starch so they are excellent energy-rich foods for humans (Shewry, 2009, Åman et al., 1985) and have a protein content of 10-15%.

1.4.1 The anatomy of the grain

The main structural components of the cereal grain are the starchy endosperm and the embryo. These components are surrounded by the aleurone layer and pericarp (the outermost part). The general anatomical structure of a grain is illustrated in figure 13.

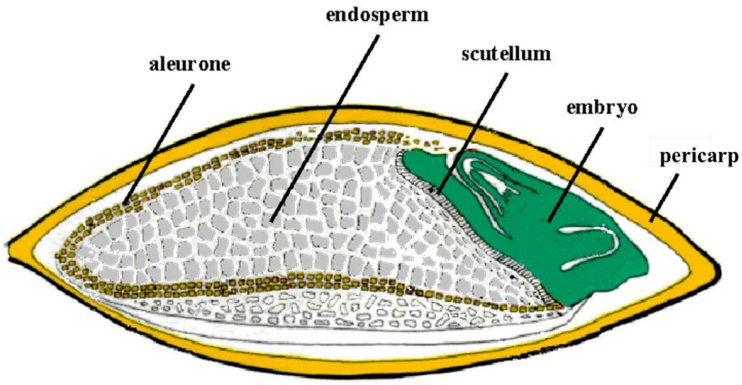


Figure 13. Main structural components of cereal grains (http://www.crc.dk/flab/the.htm).

The embryo, also known as the germ, contains various proteins, oils, enzymes and vitamins. It will develop to a mature plant when the seed has germinated. Many enzymes are synthesized during germination in the scutellum which is a part of the embryo.

The endosperm is the largest part of a grain and serves as a major storage of starch and proteins. It contains the granular starch which is trapped in a protein matrix. Starch is the energy source of the seed to be mobilized during germination. Surrounding the endosperm is **the aleurone layer** which makes up a relatively small part of the seed. During germination, the embryo produces the plant hormone gibberellic acid (GA) which triggers expression of α-amylase in the aleurone layer. The enzyme is released into the endosperm and hydrolyses starch into maltose and maltooligosaccharides which are transported to the embryo. The carbohydrate fuels respiration in the embryo so the plantlet can grow. **The pericarp** forms the outermost layer protecting the grain during development.

1.4.2 Cereal grain proteins

Cereal proteins are classically classified according to appearance in the Osborne fractionation (Osborne, 1907) which is based on the solubility properties. The Osborne fractionation divides proteins into four groups

i. Water-soluble albumins

ii. Salt-soluble globulins

iii. Alcohol-soluble gliadin (monomeric prolamins)

iv. Acid or alkaline-soluble glutelins (polymeric inter-chain disulfide bonded prolamins)

Albumins and globulins are mostly biologically-active enzymes performing catalytic and regulating functions. Most of the biologically active proteins are distributed in the compartments of aleurone layer and embryo (Lasztity et al., 1996). Gliadins and glutenins, also called prolamins play roles primarily as storage proteins (Waga, 2004). Many glutelins are structurally closely related to gliadins but form high molecular weight polymers through inter-chain disulphide bonds and are not soluble in aqueous alcohol. Prolamins have high contents of proline and glutamine. These two amino acids account for 30-70% of the total amino acids in individual prolamin proteins (Shewry, 1996). Prolamin in barley and wheat is called hordein and gluten proteins respectively. Globulins also play roles as seed storage proteins but are present in a smaller amount (Shewry and Casey, 1999, Shewry, 1996).

1.4.3 Glycosylation and applications of glycosylation analysis in cereal proteins

Some wheat and barley proteins have been reported to be glycosylated but the occupied glycosylation sites in the vast majority of these glycoproteins have not yet been identified and characterized. In addition, most information on glycosylation in cereals has been obtained from studies on individual purified glycoproteins or products of heterologous gene expression in other organisms, thus no global analysis of glycosylation sites in cereal proteins has been described. The information on identified glycosylated proteins and their glycosylation sites could provide fundamental data for further research to improve insight into wheat flour components, structure, properties, and biological roles of plant proteins including food glycoprotein allergens. In addition, such analysis may also discover novel

targets for improvement of food products from cereals, e.g. bread and beer and for the development of novel functional foods.

The allergenicity of glycoproteins is very common in plants, because the presence of xylose and the core-3-linked fucose, which do not exist in mammals. Major food allergens are glycoproteins (Lack, 2008) and the majority of IgE-mediated allergic reactions involves water soluble proteins (the albumin and globulin fractions) (Lehrer et al., 1997). Several wheat allergens from the water-soluble fraction have recently been identified by immunoblotting with IgE from patients suffering from wheat allergy towards the albumin and globulin fractions (Larré et al., 2011) and patients suffering with atopic dermatitis (Sotkovsky et al., 2011).

Barley and wheat are used in the brewing industry and many proteins present in beer are derived mainly from the barley/wheat grains. Understanding of the effects of the proteins on beer quality (e.g. foam and haze formation) requires information about characteristics and compositions on barley/wheat grain proteins. Glycated and glycosylated proteins are also reported to influence the quality of beer (Lastovickova et al., 2012), but the role of glycoproteins has not yet been elucidated in this aspect. Large content of hydroxyl group in glycosylated proteins may cause haze in beer (Leiper, 2003a,b). However, there are few publications which studied wheat and barley glycoproteins (Lastovickova et al., 2012). The list of glycoproteins in wheat and barley may provide targets for further studies to gain more insight into the effects of proteins in beer quality. In addition, the study of allergenic glycoproteins could be the target for the production of low allergenic beer.

1.5 The aim of present investigation and experimental work

N-glycosylation has been proven to be involved in several biological roles in plant such as enzyme activity (Kimura et al., 1999, Lige et al., 2001), thermostability and folding (Lige et al., 2001), oligomerization (Kimura et al., 1999), host-pathogen interaction (Lim et al., 2009) or subcellular localization and secretion (Ceriotti et al., 1998). In addition, many glycoproteins were found to be involved in allergenicity in food (Garcia-Casado et al., 1996). However, available information on glycosylation in cereal proteins (e.g. wheat and barley) is very limited and no global analysis of glycosylation sites in cereal proteins has

been described. The aims of this study were to develop a HILIC-based glycopeptide enrichment strategy and apply in a strategy for global glycoprotein identification and glycosylation site assignment in wheat albumins and proteins from barley aleurone layers.

The experimental work carried out has been described in the following chapters.

Chapter 2: Selective Glycopeptide Enrichment Using a Combination of ZIC-HILIC and Cotton Wool for Exploring the Subglycoproteome of Wheat Flour Albumins

- Describes the development of a HILIC-based glycopeptide enrichment strategy by supplement of cotton wool with ZIC-HILIC (ZIC-cotton) to provide higher selectivity of glycopeptide enrichment. By the enrichment with ZIC-cotton and cotton wool alone, 78 different glycosylation sites in 66 albumin proteins extracted from wheat flour were identified. In addition, putative wheat allergens were also predicted from these identified glycoproteins.

Chapter 3: Exploring sub *N*-glycoproteome of Barley aleurone layer

- ZIC-cotton was applied for enrichment of glycopeptides from intracellular and extracellular protein fraction of barley aleurone layers. Fourty nine glycoproteins were identified and fourteen putative allergens from barley were also predicted. Some interesting information obtained from the identified barley alerurone layer glycoproteins will be described.

Chapter 4: GA_3-induced aleurone layers responding to heat shock or tunicamycin provide insight into the *N*-glycoproteome, protein secretion and ER stress

- The effects of tunicamycin treatment and heat shock on barley aleurone layer were investigated. Because of the role of tunicamycin in blocking protein *N*-glycosylation, *N*-glycosylation proteins profile of barley aleurone layer (from chapter 3) provides potential targets for tunicamycin treatment. By 2D gel electrophoresis, many proteins corresponding to the identified glycoproteins produced from barley aleurone layer are found to decrease with either tunicamycin treatment or heat shock.

1.6 References

Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198–207.

Afrough, B.; Dwek, M. V.; Greenwel, P. Identification and elimination of false-positives in an ELISA-based system for qualitative assessment of glycoconjugate binding using a selection of plant lectins. *BioTechniques* **2007**, *43*, 458-464

Alpert, A. J. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J. Chromatogr.* **1990**, *499*, 177−196.

Alvarez-Manilla, G.; Atwood, J. 3rd.; Guo, Y.; Warren, N. L.; Orlando, R.; Pierce, M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with *N*-linked glycosylation sites. *J. proteome Res.* **2006**, *5*, 701−708.

Åman, P.; Hesselman, K.; Tilly, A. C. The variation in chemical composition of Swedish barleys. *J. Cereal Sci.* **1985**, *3*, 73−77.

An, H. J.; Froehlich, J. W.; Lebrilla, C. B. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr. Opin. Chem. Biol.* **2009**, *13*, 421−426.

Angel, P. M.; Lim, J.; Wells, L.; Bergmann, C.; Orlando, R. A potential pitfall in 18O-based *N*-linked glycosylation site mapping. *Rapid Commun. Mass Spectrom*. **2007**, *21*, 674–682.

Apweiler, R.; Hermjakob, H.; Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta.* **1999***, 1473*, 4–8.

Banerjee, S.; Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int. J. Anal. Chem.* **2012**, 1−40.

Boersema, P. J.; Mohammed, S.; Heck, A. J. Hydrophilic interaction liquid chromatography (HILIC) in proteomics. *Anal. Bioanal. Chem.* **2008**, *391*, 151−159.

Bardor, M.; Faveeuw, C.; Fitchette, A. C.; Gilbert, D.; Galas, L.; Trottein, F.; Faye, L.; Lerouge, P. Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology* **2003**, *13*, 427–434.

Beck, F.; Lewandrowski, U.; Wiltfang, M.; Feldmann, I.; Geiger, J.; Sickmann, A.; Zahedi, R. P. The good, the bad, the ugly: validating the mass spectrometric analysis of modified peptides. *Proteomics* **2011**, *11*, 1099–1109.

Berven, F. S.; Ahmad, R.; Clauser, K. R.; Carr, S. A. Optimizing performance of glycopeptide capture for plasma proteomics. *J. Proteome Res*. **2010**, *9*, 1706–1715.

Boisson, M.; Gomord, V.; Audran, C.; Berger, N.; Dubreucq, B.; Granier, F.; Lerouge, P.; Faye, L.; Caboche, M.; Lepiniec, L. *Arabidopsis* glucosidase I mutants reveal a critical role of *N*-glycan trimming in seed development. *EMBO J.* **2001**, *20*, 1010–1019.

Bunkenborg J.; Pilch B. J.; Podtelejnikov A. V.; Wiśniewski, J. R. Screening for ***N***-glycosylated proteins by liquid chromatography mass spectrometry. *Proteomics* **2004**, *4*, 454–465.

Burn, J. E.; Hocart, C. H.; Birch, R. J.; Cork, A. C.; Williamson, R. E. Functional analysis of the cellulose synthase genes CesA1, CesA2, and CesA3 in *Arabidopsis*. *Plant Physiol.* **2002**, *129*, 797–807.

Buszewski, B.; Noga, S. Hydrophilic interaction liquid chromatography (HILIC)− a powerful separation technique. *Anal. Bioanal. Chem.* **2012**, *402*, 231–247.

Catala, C.; Howe, K. J.; Hucko, S.; Rose, J. K.; Thannhauser, T. W. Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* **2011**, *11*, 1530–1544.

Ceriotti, A.; Duranti, M.; Bollini, R.; Vegetali, I. B.; Milano, I. Effects of *N*-glycosylation on the folding and structure of plant proteins. *J. Exp. Bot.* **1998**, *49*, 109–1103.

Chen, M.; Lu, Y.; Ma, Q.; Guo, L.; Feng, Y.-Q. Boronate affinity monolith for highly selective enrichment of glycopeptides and glycoproteins. *Analyst*. **2009**, *134*, 2158–2164.

Ding, W.; Nothaft, H.; Szymanski, C. M.; Kelly, J. Identification and quantification of glycoproteins using ion-pairing normal-phase liquid chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2009***, 8*, 2170–2185.

Fan, X.; She, Y. M.; Bagshaw, R. D.; Callahan, J. W.; Schachter, H.; Mahuran, D. J. A method for proteomic identification of membrane-bound proteins containing Asn-linked oligosaccharides. *Anal. Biochem.* **2004**, *332*, 178–186.

Faye, L.; Boulaflous, A.; Benchabane, M.; Gomord, V.; Michaud, D. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* **2005**, *23*, 1770–1778.

Fitchette-Laine, A. C.; Dinh, O. T.; Faye, L.; Bardor, M. Plant proteomics and glycosylation. *Methods Mol. Biol.* **2007**, *355*, 317–342.

Fitchette-Laine, A. C.; Gomord, V.; Cabanes, M.; Michalski, J. C.; Saint Macary M.; Foucher B.; Cavelier B.; Hawes, C.; Lerouge, P.; Faye, L. *N*-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J.* **1997**, *12*, 1411–1417.

Gallois, P.; Makishima, T.; Hecht, V.; Despres, B.; Laudie, M.; Nishimoto, T.; Cooke, R. An *Arabidopsis thaliana* cDNA complementing a hamster apoptosis suppressor mutant. *Plant J.* **1997**, *11*, 1325–1331.

Grinna, L. S.; Robbins, P. W. Glycoprotein biosynthesis rat liver microsomal glucosidases which process oligosaccharides. *J. Biol. Chem.* **1979**, *254*, 8814–8818.

Guo, Y.; Gaiki, S. Retention behavior of small polar compounds on polar stationary phases in hydrophilic interaction chromatography. *J. Chromatogr.* **2005**, *1074*, 71−80.

Hagglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; Roepstorff, P. A new strategy for identification of *N*-Glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res.* **2004**, *3*, 556–566.

Hampton, R. Y. ER-associated degradation in protein quality control and cellular regulation. *Curr. Opin. Cell Biol*. **2002**, *14*, 476–482.

Hao, P.; Ren, Y.; Alpert, A. J.; Sze, S. K. Detection,evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation. *Mol. Cell. Proteomics* **2011**, *10,* O111.009381.

Harvey, D. J. Matrix-assisted laser desorption/ionization mass spectrometry of carbonhydrates and glycoconjugates. *Int. J. Mass Spectrom*. **2003**, *226*, 1–35.

Helenius, A.; Aebi, M. Roles of *N*-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem*. **2004**, *73*, 1019–1049.

Hemström, P.; Irgum, K. Hydrophilic interaction chromatography. *J. Sep. Sci.* **2006**, *29*, 1784−1821.

Hogan, J. M.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. Complementary structural information from a tryptic *N*-linked glycopeptide via electron transfer ion/ion reactions and collision-induced dissociation. *J. Proteome Res.* **2005**, *4*, 628–632.

Huang, H.; Guo, H.; Xue, M.; Liu, Y.; Yang, J.; Liang, X.; Chu, C. Click novel glycosyl amino acid hydrophilic interaction chromatography stationary phase and its application in enrichment of glycopeptides. *Talanta* **2011**, *85*, 1642–1647.

Huang, H.; Jin, Y.; Xue, M.; Yu, L.; Fu, Q.; Ke, Y.; Chu, C.; Liang, X. A novel click chitooligosaccharide for hydrophilic interaction liquid chromatography. *Chem. Commun.* **2009**, *45*, 6973–6975.

Hubbard, S. C.; Ivatt, R. J. Synthesis and processing of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **1981**, *50*, 555–583.

Kaji, H.; Yamauchi, Y.; Takahashi, N.; Isobe, T. Mass spectrometric identification of *N*-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. *Nat. Protoc.* **2006**, *6*, 3019–3027.

Kaji, H.; Saito, H.; Yamauchi, Y.; Shinkawa, T.; Taoka, M.; Hirabayashi, J.; Kasai, K.-I.; Takahashi, N.; Isobe, T. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify *N*-linked glycoproteins. *Nat. Biotechnol*. **2003**, *21*, 667–672.

Kang, J. S.; Frank, J.; Kang, C. H.; Kajiura, H.; Vikram, M.; Ueda, A.; Kim, S.; Bahk, J. D.; Triplett, B.; Fujiyama, K.; Lee, S. Y.; von Schaewen, A.; Koiwa, H. Salt tolerance of *Arabidopsis thaliana* requires maturation of *N*-glycosylated proteins in the Golgi apparatus. *Proc. Natl. Acad. Sci.* **2008**, *105*, 5933–5938.

Kaufmann, R.; Spengler, B.; Lutzenkirchen, F. Mass-spectrometric sequencing of linear peptides by product-ion analysis in a reflectron time-of-flight mass-spectrometer using matrix-assisted laser- desorption ionization. *Rapid Commun. Mass Spectrom*. **1993**, *7*, 902–910.

Kaufmann, R.; Kirsch, D.; Spengler, B. Sequencing of peptides in a time-of-flight mass-spectrometer - evaluation of postsource decay following matrix-assisted laser-desorption ionization (MALDI). *Int. J. Mass Spectrom. Ion Process* **1994**, *131*, 355–385.

Kimura, Y.; Hess, D.; Sturm, A. The *N*-glycans of jack bean α-mannosidase. Structure, topology and function. *Eur. J. Biochem.* **1999**, *264,* 168–175.

Klement, E.; Udvardy, A.; Medzihradszky, K. F. Enrichment of O-GlcNAc modified proteins by the periodate oxidation-hydrazide resin capture approach. *J. Proteome Res.* **2010**, *9*, 2200–2206.

Knauer, R.; Lehle, L. The oligosaccharyl transferase complex from *Saccharomyces cerevisiae*. Isolation of the OST6 gene,its synthetic interaction with OST3, and analysis of the native complex. *J. Biol. Chem.* **1999**, *274*, 17249–17256.

Koide, N.; Muramatsu, T. J. Endo-β-N-acetylglucosaminidase acting on carbohydrate moieties of glycoproteins. Purification and Properites of The enzyme from *Diplococcus pneumoniae*. *Biol. Chem.* **1974**, 249, 4897–4904.

Koiwa, H.; Li, F.; Mccully, M. G.; Mendoza, I.; Koizumi, N.; Manabe, Y.; Nakagawa, Y.; Zhu, J.; Rus, A.; Pardo, J. M.; Bressan, R. A.; Hasegawa, P. M. The STT3 a subunit isoform of the *Arabidopsis* oligosaccharyltransferase controls adaptive responses to salt osmotic stress. *Plant Cell* **2003**, *15*, 2273–2284.

Kuo, C. W.; Wu, I. L.; Hsiao, H. H.; Khoo, K. H. Rapid glycopeptide enrichment and N-glycosylation site mapping strategies based on amine-functionalized magnetic nanoparticles. *Anal. Bioanal. Chem.* **2012**, *402*, 2765–2776.

Lack, G. Clinical practice: food allergy. *N. Engl. J. Med*. **2008**, *359*, 1252–1260.

Larré, C.; Lupi, R.; Gombaud, G.; Brossard, C.; Branlard, G.; Moneret-Vautrin, D. A.; Rogniaux, H. Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction. *J. Proteomics* **2011**, *74*, 1279–1289.

Laštovičková, M.; Bobálová, J. MS based proteomic approaches for analysis of barley malt. *J. Cereal Sci.* **2012**, *56*, 519–530.

Lásztity, R. The importance and general characterization of cereal proteins. In *The Chemistry of Cereal Proteins*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 1996; pp. 3−17.

Lee, A.; Kolarich, D.; Haynes, P. A.; Jensen, P. H.; Baker, M. S.; Packer, N. H. Rat liver membrane glycoproteome: enrichment by phase partitioning and glycoprotein capture. *J. Proteome Res.* **2009**, *8*, 770–781.

Lehrer, S. B.; Taylor, S. L.; Helfe, S. L.; Bush, R. K. Food allergens. In *Allergy and allergic diseases*, 1st ed.; Kay, A. B., Ed.; Blackwell Sciences: London, UK, 1997; pp. 961−980.

Leiper, K. A.; Stewart, G. G.; McKeown, I. P. Beer polypeptides and silica gel-Part I. Polypeptides involved in haze formation. *J. Inst. Brew*. **2003**a, *109*, 57–72.

Leiper, K. A.; Stewart, G. G.; McKeown, I. P. Beer polypeptides and silica gel-Part II. Polypeptides involved in foam formation. *J. Inst. Brew*. **2003**b, *109*, 73–79.

Lerouge, P.; Cabanes-Macheteau, M.; Rayon, C.; Fischette-Lainé, A. C.; Gomord, V.; Faye, L. *N*-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant mol. Biol.* **1998**, *38*, 31–48.

Lerouxel, O.; Mouille, G.; Andeme-Onzighi, C.; Bruyant, M. P.; Seveno, M.; Loutelier-Bourhis, C.; Driouich, A.; Hofte, H.; Lerouge, P. Mutants in defective glycosylation, an *Arabidopsis* homolog of an oligosaccharyl transferase complex subunit show protein under glycosylation and defects in cell differentiation and growth. *Plant J.* **2005**, *42*, 455–468.

Lewandrowski, U.; Zahedi, R. P.; Moebius, J.; Walter, U.; Sickmann, A. Enhanced *N*-glycosylation site analysis of sialoglycopeptides by strong cation exchange prefractionation applied to platelet plasma membranes. *Mol. Cell Proteomics* **2007**, *6*, 1933−1941.

Li, L.; Wang, L.; Zhang, W.; Tang, B.; Zhang, J.; Song, H.; Yao, D.; Tang, Y.; Chen, X.; Yang, Z.; Wang, G.; Li, X.; Zhao, J.; Ding, H.; Reed, E.; Li, Q. Q. Correlation of serum VEGF levels with clinical stage, therapy efficacy, tumor metastasis and patient survival in ovarian cancer. *Anticancer Res.* **2004**, *24*, 1973–1979.

Liebminger, E.; Veit, C.; Mach, L.; Strasser, R. Mannose trimming reactions in the early stages of the *N*-glycan processing pathway. *Plant Signal Behav.* **2010**, *5*, 476–478.

Lige, B.; Ma, S.; van Huystee, R. B. The effects of the site-directed removal of *N*-glycosylation from cationic peanut peroxidase on its function. *Arch. Biochem. Biophys.* **2001**, *386*, 17–24.

Linden, J. C.; Lawhead, C. L. Liquid chromatography of saccharides. *J. Chromatogr., A* **1975**,105, 125-133.

Lim, J. M.; Aoki, K.; Angel P.; Garrison, D.; King, D.; Tiemeyer, M.; Bergmann, C.; Wells, L. Mapping glycans onto specific *N*-linked glycosylation sites of *Pyrus communis* PGIP redefines the interface for EPG-PGIP interactions. *J. Proteome Res.* **2009**, *8*, 673–680.

Mechref, Y. Use of CID/ETD mass spectrometry to analyze glycopeptides. *Curr. Protoc. Protein Sci.* **2012**, *12*, 1-11.

Melo, N. S.; Nimtz, M.; Conradt, H. S.; Fevereiro, P. S.; Costa, J. Identification of the human Lewisa carbohydrate motif in a secretory peroxidase from a plant cell suspension culture (*Vaccinium myrtillus* L.). *FEBS Lett*. **1997**, *415*, 186–191.

Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R.A. Localization of *O*-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal.Chem.* **1999**, *71*, 4431–4436.

Minic Z.; Jamet, E.; Négroni, L.; der Garabedian, P. A.; Zivy, M.; Jouanin, L. A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J. Exp. Bot.* **2007**, *58*, 2503–2512.

Mysling, S.; Palmisano, G.; Højrup, P.; Thaysen-Andersen, M. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. *Anal. Chem*. **2010**, *82*, 5598–5609.

Ohyama, Y.; Kasai, K.; Nomoto, H.; Inoue, Y. Frontal affinity chromatography of ovalbumin glycoasparagines on a concanavalin A-sepharose column. A quantitative study of the binding specificity of the lectin. *J. Biol. Chem.* **1985**, *260*, 6882–6887.

Ongay, S.; Boichenko, A.; Govorukhina, N.; Bischoff, R. Glycopeptide enrichment and separation for protein glycosylation analysis. *J. Sep. Sci.* **2012**, *35*, 2341–2372.

Osborne, T. B. The proteins of the wheat kernel. In *The Carnegie Institution of Washington*, Carnegie Institute: Washington, DC, USA, 1907; Publ 84, pp. 1−119.

Palmisano, G.; Antonacci, D.; Larsen, M. R. Glycoproteomic profile in wine: a " sweet " molecular renaissance research. *J. Proteome Res*. **2010**, *9*, 6148–6159.

Parker, B. L.; Palmisano, G.; Edwards, A. V. G.; White, M. Y.; Engholm-Keller, K.; Lee, A.; Scott, N. E.; Kolarich, D.; Hambly, B. D.; Packer, N. H.; Larsen, M. R.; Cordwell, S. J. Quantitative *N*-linked glycoproteomics of myocardial ischemia and reperfusion injury reveals early remodeling in the extracellular environment. *Mol. Cell. Proteomics* **2011**, *10,* M110.006833.

Pasing, Y.; Sickmann, A.; Lewandrowski, U. *N*-glycoproteomics: mass spectrometry-based glycosylation site annotation. *Biol. Chem.* **2012**, *393*, 249–258.

Patton, W. F. Detection technologies in proteome analysis. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **2002**, *771*, 3-31.

Petrescu, A.-J.; Milac, A.-L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. Statistical analysis of the protein environment of *N*-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* **2004**, *14*, 103–114.

Pless, D. D.; Lennarz, W. J. Enzymatic conversion of proteins to glycoproteins. *Proc. Natl. Acad. Sci.* **1977**, *74*, 134–138.

Plummer, T. H.; Tarentino, A. L. Facile cleavage of complex oligosaccharides from glycopeptides by almond emulsin peptide: *N*-glycosidase. *J. Biol. Chem.* **1981**, *256*, 10243–10246.

Pompach, P.; Chandler, K. B.; Lan, R.; Edwards, N.; Goldman, R. Semi-automated identification of *N*-Glycopeptides by hydrophilic interaction chromatography, nano-reverse-phase LC-MS/MS, and glycan database search. *J. Proteome Res.* **2012**, *11*, 1728–1740.

Qi, D.; Zhang, H.; Tang, J.; Deng, C.; Zhang, X. Facile synthesis of mercaptophenylboronic acid-functionalized core-shell structure Fe_3O_4@C@Au magnetic microspheres for selective enrichment of glycopeptides and glycoproteins. *J. Phys. Chem. C.* **2010**, *114*, 9221–9226.

Roepstorff, P.; Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **1984**, *11*, 601.

Ruiz-May, E.; Thannhauser, T. W.; Zhang, S.; Rose, J. K. C. Analytical technologies for identification and characterization of the plant *N*-glycoproteome. *Front. Plant Sci.* **2012**, *3*, 1-8.

Schollen, E.; Grunewald, S.; Keldermans, L.; Albrecht, B.; Korner, C.; Matthijs, G. CDG-Id caused by homozygosity for an ALG3 mutation due to segmental maternal isodisomy UPD3(q21.3-qter). *Eur. J. Med. Genet.* **2005**, *48*, 153–158.

Scott, N. E.; Parker, B. L.; Connolly, A. M.; Paulech, J.; Edwards, A. V.; Crossett, B.; Falconer, L.; Kolarich, D.; Djordjevic, S. P.; Hojrup, P.; Packer, N. H.; Larsen, M. R.; Cordwell, S. J. Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, HCD and ETD-MS applied to the *N*-

linked glycoproteome of Campylobacter jejuni. *Mol. Cell. Proteomics* **2011,** *10,* M000031-MCP201.

Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M.; Wuhrer, M. Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal. Chem.* **2011**, *83*, 2492–2499.

Sharon, N. Lectin-carbohydrate complexes of plants and animals: an atomic view. *Trends Biochem. Sci*. **1993**, *18*, 221–226.

Shewry, P. R. Cereal grain proteins. In *The Cereal grain quality*, Henry, R. J., Kettlewell, P. S., Eds.; Chapman and Hall: London, UK, 1996; pp. 227–250.

Shewry, P. R. Wheat. *J. Exp. Bot.* **2009**, *60*, 1537–1553.

Shewry, P. R.; Casey, R. Seed proteins. In *The Seed Proteins*, Shewry, P. R., Casey, R., Eds.; Kluwer Academic Publishers: Dordrecht, NL, 1999; pp. 1–10.

Song, W.; Henquet, M. G.; Mentink, R. A.; van Dijk, A. J.; Cordewener, J. H.; Bosch, D.; America, A. H.; van der Krol A. R. *N*-glycoproteomics in plants: perspectives and challenges. *J. Proteomics* **2011**, *74*, 1463–1474.

Sojar, H. T.; Bahl, O. P. A chemical method for the deglycosylation of proteins. Arch. Biochem. Biophys. **1987**, *259*, 52–57.

Sotkovsky, P. S.; Sklenar, J.; Halada, P.; Cinova, J.; Etinova, I. S.; Kainarova, A.; Golias, J.; Pavlaskova, K.; Honzova S.; Tuckova, L. A new approach to the isolation and characterization of wheat flour allergens. *Clin. Exp. Allergy* **2011**, *41*, 1031–1043.

Spiro, R. G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **2002**, *12*, 43R–56R.

Steen, H.; Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol*. **2004**, *5*, 699−711.

Strasser, R.; Stadlmann, J.; Schahs, M.; Stiegler, G.; Quendler, H.; Mach, L.; Glössl, J.; Weterings, K.; Pabst, M.; Steinkellner, H. Generation of glyco-engineered *Nicotiana*

benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* **2008**, *6*, 392–402.

Strasser, R.; Stadlmann, J.; Svoboda, B.; Altmann, F.; Glössl, J.; Mach, L. Molecular basis of N-acetylglucosaminyltransferase I deficiency in *Arabidopsis thaliana* plants lacking complex *N*-glycans. *Biochem. J.* **2005**, *387*, 385–391.

Stroud, D.A.; Oeljeklaus, S.; Wiese, S.; Bohnert, M.; Lewandrowski, U.; Sickmann, A.; Guiard, B.; van der Laan, M.; Warscheid, B.; Wiedemann, N. Composition and topology of the endoplasmic reticulum-mitochondria encounter structure. *J. Mol. Biol.* **2001**, *413*, 743–750.

Sun, B.; Ranish, J. A.; Utleg, A. G.; White, J. T.; Yan, X.; Lin, B.; Hood, L. Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. *Mol. Cell. Proteomics* **2007***, 6*, 141–149.

Swaney, D. L.; McAlister, G. C.; Wirtala, M.; Schwartz, J. C.; Syka, J. E.; Coon, J. J. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal. Chem.* **2007**, *79*, 477−485.

Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci.* **2004**, *101*, 9528–9533.

Szumilo, T.; Kaushal, G. P.; Elbein, A. D. Demonstration of GlcNAc transferase I in plants. *Biochem. Biophys. Res. Commun.* **1986**a, *134*, 1395–1403.

Szumilo, T.; Kaushal, G. P.; Hori, H.; Elbein, A. D. Purification and properties of a glycoprotein processing α-mannosidase from mung bean seedlings. *Plant Physiol.* **1986**b, *81*, 383–389.

Taylor, M. A.; Ross, H. A.; Mcrae, D.; Stewart, D.; Roberts, I.; Duncan, G.; Wright, F.; Millam, S.; Davies, H. V. A potato α-glucosidase gene encodes a glycoprotein processing α-glucosidase II-like activity. Demonstration of enzyme activity and effects of down-regulation in transgenic plants. *Plant J.* **2000**, *24*, 305–316.

Tretter, V.; Altmann, F.; März, L. Peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidase F cannot release glycans with fucose attached alpha 1→3 to the asparagine-linked *N*-acetylglucosamine residue. *Eur. J. Biochem.* **1991**, *199*, 647–652.

Ueda, K.; Takami, S.; Saichi, N.; Daigo, Y.; Ishikawa, N.; Kohno, N.; Katsumata, M.; Yamane, A.; Ota, M.; Sato, T. A.; Nakamura, Y.; Nakagawa, H. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol. Cell. Proteomics* **2010**, *9*, 1819–1828.

von Schaewen, A.; Sturm, A.; O'Neill, J.; Chrispeels, M. J. Isolation of a mutant *Arabidopsis* plant that lacks *N*-acetyl glucosaminyltransferase I and is unable to synthesize Golgi-modified complex *N*-linked glycans. *Plant Physiol*. **1993**, *102*, 1109–1118.

Wada, Y.; Tajiri, M.; Yoshida, S. Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem*. **2004**, *76*, 6560–6565.

Waga, J. Structure and allergenicity of wheat gluten proteins – a review. *J. Food Nutr. Sci.* **2004***, 13*, 327−338.

Wan, H.; Yan, J.; Yu, L.; Sheng, Q.; Zhang, X.; Xue, X.; Li, X.; Liang, X. Zirconia layer coated mesoporous silica microspheres as HILIC SPE materials for selective glycopeptide enrichment. *Analyst.* **2011**, *136*, 4422−4430.

Wiesner, J.; Premsler, T.; Sickmann, A. Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. *Proteomics* **2008**, *8*, 4466–4483.

Wohlgemuth, J.; Karas, M.; Eichhorn, T.; Hendriks, R.; Andrecht, S. Quantitative site-specific analysis of protein glycosylation by LC-MS using different glycopeptide-enrichment strategies. *Anal. Biochem.* **2009**, *395*, 178–188.

Wuhrer, M.; Catalina, M. I.; Deelder, A. M.; Hokke, C. H. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2007**, *849*, 115–128.

Wuhrer, M.; de Boer, A. R.; Deelder, A. M. Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom. Rev.* **2009**, *28*, 192–206.

Xu, Y.; Wu, Z.; Zhang, L.; Lu, H.; Yang, P.; Webley, P. A.; Zhao, D. Highly specific enrichment of glycopeptides using boronic acid-functionalized mesoporous silica. *Anal. Chem*. **2009**, *81*, 503–508.

Yamamoto, K.; Tsuji, T.; Osawa, T. Analysis of asparagine-linked oligosaccharides by sequential lectin affinity chromatography. *Mol. Biotechnol*. **1995**, *3*, 25–36.

Yamamoto,K.; Tsuji,T.; Osawa, T. Analysisofasparagine- linked oligosaccharides by sequential lectin-affinity chromatography. *Methods Mol. Biol.***1998**, *76*, 35–51.

Yan, Q.; Lennarz, W. J. Oligosaccharyl transferase: a complex multisubunit enzyme of the endoplasmic reticulum. *Biochem. Biophys. Res.Commun.* **1999**, *266*, 684–689

Yang, Z.; Harris, L. E.; Palmer-Toy, D. E.; Hancock W. S. Multilectin affinity chromatography for characterization of multiple glycoprotein biomarker candidates in serum from breast cancer patients. *Clin. Chem.* **2006**, *52*, 1897–1905.

Yeh, C. H. Chen, S. H.; Li, D. T.; Lin, H. P.; Huang, H. J.; Chang, C. I.; Shih, W. L.; Chern, C. L.; Shi, F. K.; Hsu, J. L. Magnetic bead-based hydrophilic interaction liquid chromatography for glycopeptide enrichments. *J. Chromatogr. , A.* **2012**, *10*, 70-78.

Yu, L.; Li, X.; Guo, Z.; Zhang, X.; Liang, X. Hydrophilic interaction chromatography based enrichment of glycopeptides by using click maltose: a matrix with high selectivity and glycosylation heterogeneity coverage. *Chemistry* **2009**, *15*, 12618−12626.

Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **2003**, *21*, 660–666.

Zhang, H.; Yao, G.; Deng, C.; Lu, H.; Yang, P. Facile synthesis of boronic acid-functionalized magnetic mesoporous silica nanocomposites for highly specific enrichment of glycopeptides. *Chin. J. Chem.* **2011**, *29*, 835–839.

Zhang, Q.; Schepmoes, A. A.; Brock, J. W. C.; Wu, S.; Moore, R. J.; Purvine, S. O.; Baynes, J. W.; Smith R. D.; Metz T. O. Improved methods for the enrichment and analysis of glycated peptides. *Anal. Chem.* **2008**, *80*, 9822–9829.

Zhang, Y.; Giboulot, A.; Zivy, M.; Valot, B.; Jamet, E.; Albenne, C. Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* **2010**, *72*, 1109–1123.

Zhao, Y.; Yu, L.; Guo, Z.; Li, X.; Liang, X. Reversed-phase depletion coupled with hydrophilic affinity enrichment for the selective isolation of *N*-linked glycopeptides by using Click OEG-CD matrix. *Anal. Bioanal. Chem.* **2011**, *399*, 3359–3365.

Zhou, Y.; Aebersold, R.; Zhang, H. Isolation of *N*-linked glycopeptides from plasma. *Anal. Chem.* **2007**, *79*, 5826–5837.

Zielinska, D. F.; Gnad, F.; Wiśniewski, J. R.; Mann, M. Precision mapping of an in vivo *N*-glycoproteome reveals rigid topological and sequence constraints. *Cell* **2010**, *141*, 897–907.

Chapter 2 — Glycopeptide Enrichment Using a Combination of ZIC-HILIC and Cotton Wool for Exploring the Subglycoproteome of Wheat Flour Albumins

Glycopeptide Enrichment Using a Combination of ZIC-HILIC and Cotton Wool for Exploring the Glycoproteome of Wheat Flour Albumins

Plaipol Dedvisitsakul[1], Susanne Jacobsen[1], Birte Svensson[1], Jakob Bunkenborg[2], Christine Finnie[3], and Per Hägglund[1,*]

[1] Enzyme and Protein Chemistry, Søltofts Plads Building 224, Department of Systems Biology, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark

[2] Department of Clinical Biochemistry, Copenhagen University Hospital Hvidovre, DK-2650 Hvidovre Denmark

[3] Agricultural and Environmental Proteomics, Søltofts Plads Building 224, Department of Systems Biology, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark

*Corresponding author: Per Hägglund, Enzyme and Protein Chemistry, Søltofts Plads Building 224, Department of Systems Biology, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark

Tel: +45 4525 5503; fax: +45 4588 6307; e-mail: ph@bio.dtu.dk

ABSTRACT

Hydrophilic liquid chromatography (HILIC) is used extensively as a sample preparation step for glycopeptide enrichment in proteome research. Here, we have applied cotton wool and a zwitterionic HILIC (ZIC-HILIC) resin to provide higher loading capacity and broader specificity for glycopeptide enrichment. This strategy was applied to wheat flour albumin extracts followed by site-specific glycosylation labeling and mass spectrometry for assignment of 78 *N*-glycosylation sites in 67 albumin proteins. Bioinformatics analysis revealed that several of the identified glycoproteins show sequence similarity to known food allergens. In addition, the potential impact of some of the identified glycoproteins on wheat beer quality is discussed.

KEYWORDS: glycopeptide enrichment, HILIC, cotton, *N*-glycosylation, mass spectrometry, deglycosylation, deamidation, isotope labeling, cereal proteomics

INTRODUCTION

Glycosylation is a major post-translational protein modification known to affect protein folding, stability, activity, molecular interaction and functionality.[1] *N*- and *O*-linked glycosylations represent the most common types of glycosylation.[2] The majority of proteins secreted into the endoplasmic reticulum (ER) are modified by attachment of an oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) onto asparagine residues in a N-X-S/T/C consensus sequences,[3, 4] where X can be any amino acid other than proline.[5] In contrast, there appears to be no strict consensus sequence guiding the attachment of *O*-linked glycans from soluble nucleotides to the hydroxyl groups of serine or threonine. During transport along the secretory pathway, the *N*-linked oligosaccharide undergoes a number of maturation steps. First, glucose (Glc) and mannose (Man) residues are removed, and subsequently new sugar residues are added to generate either high mannose-, hybrid-, or complex-type *N*-glycans.[6] For each of these types a wide variety of glycan structures have been described and the occupancy of these glycans at different glycosylation sites lead to glycoprotein heterogeneity. Plant complex *N*-glycans differ from those produced in mammals. For instance, α(1,6)-fucose (Fuc) and terminal *N*-acetyl-neuraminic acid (NeuAc) are characteristic of mammals, whereas bisecting β(1,2)-xylose (Xyl) and core α(1,3)-Fuc are found in plants.[6] Currently very little information is available on the functional importance of glycosylated proteins from plants as compared to mammals. Some reports however on plant glycoproteins conclude that glycosylation is required for enzyme activity,[7, 8] thermostability and folding,[8] oligomerization[7] and host-pathogen interaction.[9]

In mass spectrometry ionization of glycopeptides is less efficient compared to non-glycosylated peptides leading to signal suppression and poor detection of glycopeptides[10]. Therefore, enrichment of glycopeptides prior to mass spectrometry is often required. Several approaches for glycopeptide enrichment are available based on e.g. affinity towards glycan-specific lectins[11, 12] and *cis*-diol periodate oxidation followed by hydrazide coupling.[13, 14] Glycopeptides are also frequently separated from more hydrophobic non-glycosylated peptides by hydrophilic interaction liquid chromatography (HILIC). Several HILIC materials have been used for glycopeptide enrichment, such as zwitterionic (ZIC),[15, 16, 17, 18] amide-based (e.g. TSKgel Amide-80)[18] and carbohydrate-

based HILIC material (e.g. Sepharose, cellulose and cotton wool).[19, 20] The retention of glycans in ZIC-HILIC involves polar and ionic interactions,[21] but with carbohydrate-based materials such as cotton, the retention is mainly based on hydrogen bonding.[20] These HILIC materials may thus provide distinctly different glycopeptide profiles. Not only the glycan but also the amino acid composition determines the hydrophilicity of glycopeptides. Thus ion-pairing agents neutralizing charges on amino acid residues have been demonstrated to improve glycopeptide enrichment by ZIC-HILIC.[22]

By comparison of different HILIC materials, ZIC-HILIC showed the highest selectivity and signal intensity for glycopeptides and resulted in low background of non-glycosylated peptides.[18] In addition to *N*- and *O*-linked glycopeptides, ZIC-HILIC has also successfully been applied for enrichment of glycosylphosphatidylinositol-anchored glycopeptides.[23] For proteomics analysis, the ZIC-HILIC material is typically prepared in a custom-made "micro-column" packed in a pipette tip.[24, 16] Construction of these columns is however technically challenging due to the small particle size (10 µM) and the loading capacity is limited. One way to increase capacity problem is to increase the amount of ZIC-HILIC resin in the micro-column, but flow velocity is very slow due to high backpressure. HILIC columns based on cotton wool, on the other hand, are cheap and very easy to prepare and also give fast flow properties.[20]

In the present work, we implemented a combination of cotton wool and ZIC-HILIC (referred to as ZIC-cotton HILIC) for larger scale glycopeptide enrichment in a micro-column format compatible with sensitive mass spectrometric analysis. ZIC-cotton was applied for enrichment of tryptic *N*-glycopeptides prepared from wheat flour albumin proteins followed by deglycosylation and LC-MS/MS analysis for glycosylation site mapping. Only few wheat glycoproteins have to date been identified and in most cases details on the glycosylation including types and sites remain to be described. In addition, most information on glycosylation in wheat was obtained from studies of a single purified glycoprotein or after heterologous gene expression, thus the global picture of glycosylation sites in wheat proteins is not yet described. The glycoproteins and glycosylation sites identified in the present study thus contribute to a more comprehensive description of wheat flour components and provide a basis for future research on biological function of plant proteins and food glycoprotein allergens. The long-term benefits include discovering novel targets for

improvement of wheat-based food products (e.g. bread and beer) and the development of novel functional foods.

MATERIALS AND METHODS

Protein samples

Wheat albumin proteins were prepared as previously described[25] with some modifications. Briefly, 100 mg grain powder was mixed with 25 mM sodium phosphate buffer pH 7.5 (1 mL) at 4 ˚C for 60 min, insoluble debris was removed by centrifugation and the supernatant was precipitated by an equal volume of 20% TCA. Horseradish peroxidase (HRP) and bovine serum albumin (BSA) were purchased from Sigma-Aldrich.

In-solution trypsin digestion

Protein samples (approximately 40 µg) were re-dissolved in 400 mM NH_4CO_3, 8 M urea (21 µL) and reduced by addition of 5 µL 45 mM DTT incubated at 50°C for 20 min and thereafter chilled on ice. Cysteine alkylation was performed by adding 100 mM iodoacetamide (5 µL) and incubating at room temperature (45 min). Finally, 140 µL H_2O and 5 µL (0.04 µg/mL) sequencing grade modified trypsin (Promega, Madison, WI) was added followed by incubation at 37°C for 24 h. Samples were subsequently stored at -20 °C until use.

Glycopeptide enrichment

Cotton wool and ZIC-HILIC micro-columns were prepared essentially as described.[20,16] Cotton wool micro-columns were made by pushing a small piece of cotton wool taken from a cotton wool pad into a 20 µL GELoader tip (Eppendorf, Hamburg, Germany). For ZIC-cotton HILIC micro-columns, the ZIC-HILIC resin (kind gift of Sequant, Umeå, Sweden; particle size 10 μm) was packed on top of the cotton wool (Supporting information Figure 1). In-solution trypsin digested protein samples (20 µg) were dried (SPD 1010 Speed Vac, Thermo Scientific), redissolved in either 80% (v/v) acetonitrile (ACN), 0.1% of trifluoroacetic acid (TFA) or 80% (v/v) acetonitrile (ACN), 0.5% (v/v) formic acid (FA) and applied onto HILIC micro-columns equilibrated with the same solvents. The

column was subsequently washed twice followed by elution of bound peptides by 99.5% H_2O, 0.5% FA.

Deglycosylation by PNGase A and ^{18}O-labeling

Glycopeptide eluates from ZIC-cotton HILIC micro-columns were dried (SPD 1010 Speed Vac) and re-dissolved in 100 mM ammonium acetate (pH 5.5) prepared in $H_2{}^{18}O$. Deglycosylation was performed for 24 h at 37°C with 0.1 mU *N*-glycosidase A (PNGase A, Roche Applied Science). Subsequently, deglycosylated samples were dried (SPD 1010 Speed Vac) and re-dissolved in 100 mM ammonium acetate (pH 5.5) in normal water ($H_2{}^{16}O$). Samples were added 0.2 µg of sequencing grade modified trypsin (Promega, Madison, WI) and incubated as above for 24 h at 37°C.

MALDI-TOF MS

Samples (1 µL) were spotted onto a ground steel MALDI target plate (Bruker Daltonics, Bremen Germany), allowed to dry, and overlaid with 1 µL of dihydroxybenzoic acid (DHB) matrix, 20 mg/mL in 99.5% H_2O, 0.5% FA. MALDI TOF MS analyses were performed by Ultraflex II MALDI-TOF/TOF MS (Bruker Daltonics). The obtained mass spectra were processed with FlexAnalysis software (Bruker Daltonics).

LC-MS/MS and database search

Peptide samples were desalted with StageTip C18-reversed phase columns (Thermo Scientific) and analysed on an EASY nLC 1000 chromatograph (Thermo Scientific) with an EASY-Spray column (Pepmap 3 µm, C18 15 cm x 75 µm) coupled on-line to a Q Exactive^{TM} orbitrap (Thermo Scientific). The separation was typically performed using a flow rate of 300 nL min^{-1} and a gradient of solvents A (0.1% FA) and B (80% ACN, 0.1% FA). Raw data acquired in the positive ion mode were converted into the Mascot generic format (MGF) by Proteome Discoverer 1.3 (Thermo Scientific). Data were searched against against HarvEST35 (http://www.harvest-web.org/) and TIGR wheat (http://compbio.dfci.harvard.edu/tgi/) EST databases using MASCOT (Version 2.2.04, Matrix Science). Fixed modifications: carbamidomethyl cysteine; variable modifications: oxidation of methionine, deamidation of Asn with incorporation of ^{16}O and ^{18}O; enzyme: trypsin (1 missed

cleavage). MS and MS/MS accuracy were set to 10 ppm and 20 mmu, respectively. Only glycopeptides containing the consensus tripeptide (NXS/T/C) in which the deamidated N was tagged with ^{18}O atoms were accepted. The identified EST sequences were annotated by BLAST searches against NCBI (http://www.ncbi.nlm.nih.gov). In addition, the identified protein sequences from the HarvEST35 and TIGR wheat databases were searched against the Structural Database of Allergenic Proteins (SDAP http://fermi.utmb.edu/SDAP/) for bioinformatics-based allergen predictions.

RESULTS AND DISCUSSION

Enrichment of glycopeptides by HILIC

Microcolumns containing reversed phase or ZIC-HILIC resins, cotton wool or a combination of ZIC-HILIC resin and cotton wool (ZIC-cotton HILIC) were constructed and the efficiency for enrichment of glycopeptides was tested by applying a protein model mixture of the glycoprotein horseradish peroxidase and non-glycosylated BSA (HRP-BSA). Peptides from tryptic digests of HRP-BSA were thus loaded on microcolumns packed with the various resins and analyzed by MALDI-TOF MS (Figure 1). In the sample eluted from a reversed phase microcolumn a range of high intensity signals from HRP-BSA peptides matched predicted non-glycosylated peptides (Figure 1 g). In contrast, the mass spectra of samples eluted from HILIC columns (Figure 1 a-f) are in general dominated by signals in a higher m/z range (>3000) matching previously reported glycopeptides.[26] While weak signals of a considerable number of non-glycosylated peptides were observed after enrichment by cotton wool (Figure 1 b, e), most of these were absent in the mass spectra from ZIC-cotton (Figure 1a, d). This indicates that cotton wool supplemented with ZIC-HILIC resin show increased selectivity for glycopeptides. A lower number of glycosylated peptides were detected with ZIC-HILIC, compared to cotton wool and ZIC-cotton (Figure 1 c, f). This result possibly reflects either a more narrow specificity of ZIC-HILIC compared to cotton wool or lower capacity of ZIC-HILIC due to the limited amount of resin in the microcolumn.

The use of TFA instead of FA in eluents did not efficiently reduce the binding of non-glycosylated peptides (Figure 1). In an acidic mobile phase, the negatively charged TFA ion-pairs with positively

charged amino acids thereby reducing both hydrophilic and weak electrostatic interactions of non-glycosylated peptides with the ZIC-HILIC resin.[22] Cotton wool is however uncharged and non-glycosylated peptides are expected to be retained mainly via hydrogen bond interactions that are not blocked by ion pairing with TFA. The co-enriched non-glycosylated peptides ETYGDMADCCEKQEPER (m/z 2021.1), VHKECCHGDLLECADDR (m/z 2115.3), FKDLGEEHFK (m/z 1250.4), YICDNQDTISSK (m/z 1444.5), RPCFSALTPDETYVPK (m/z 1882.1), and DDPHACYSTVFDK (m/z 1555.7), contain side-chain carboxylic acids (asparate and glutamate), hydroxyls (serine and threonine) and amides (carboxyamidomethylated cysteine) which may hydrogen bond with the abundant hydroxyl groups in the cellulose.

Deglycosylation of glycopeptides and ^{18}O labeling

HRP-BSA peptides eluted from HILIC microcolumns were subjected to deglycosylation using PNGase A that hydrolyses the *N*-glycan bond between asparagine and the innermost GlcNAc of glycans with or without α-1,3-linked core fucosylation that is common in glycoproteins from insects[27]and plants[6]. In this process the asparagine residue is converted to aspartate through deamidation. Glycosylation sites can thus be determined through the resulting net mass increase of 0.98 Da. In order to distinguish the glycosylation sites from other deamidated asparagines, deglycosylation is performed in $H_2{}^{18}O$ resulting in a unique mass increment of 2.99 Da for residues subjected to the reaction with PNGase A that may be used in database search programs.[28,29]

However, when the mass spectra obtained from non-glycosylated peptides in HRP-BSA tryptic digests subjected to PNGase A in $H_2{}^{18}O$ and $H_2{}^{16}O$ were compared, a mass shift of approximately 2 Da was observed, indicating that ^{18}O-labeling is not specific to glycosylation sites (Figure 2). This is probably due to incorporation of ^{18}O in the C-terminal carboxylate of the peptide backbone mediated through exchange catalyzed by residual trace amounts of trypsin as previously reported.[30] Labeling of C-terminal carboxyl groups complicates data analysis and may result in false positive identification of glycosylation sites.[31] To minimize ^{18}O-labeling of the C-terminus, we thus subjected the ^{18}O-labeled peptides to a second round of carboxyl-oxygen exchange activity with trypsin in the presence of $H_2{}^{16}O$. After this exchange reaction no mass shifts were observed for non-glycosylated peptides subjected to PNGase A in $H_2{}^{18}O$ and $H_2{}^{16}O$, while peptides with single and double *N*-glycosylation sites displayed peak shifts of approximately +2 Da and +4 Da,

respectively, indicative of successful incorporation of ^{18}O at the deamidated asparagines and depletion from the C-termini (Figure 2). Attempts to minimize ^{18}O-labeling at the C-termini by using a protease inhibitor (Pefabloc) and heat inactivation of trypsin were less efficient than the trypsin-catalyzed back exchange method (Supporting information Figure 2).

A strategy for global analysis of *N*-glycosylation sites in wheat protein extracts

Based on the results from the model system with HRP-BSA a strategy was devised for global screening of *N*-glycosylation sites in proteins from the albumin fraction of wheat flour (Figure 3). Firstly, cotton wool and ZIC-cotton microcolumns were employed to enrich glycopeptides from tryptic digests using mobile phases containing either 0.1% TFA or 0.5% FA in 80% ACN. To enable preliminary assessment of the enrichment efficiency, samples of eluted wheat albumin tryptic peptides were analyzed by MALDI-TOF MS as shown in Figure 4. Signals at higher m/z ratios are seen in samples enriched by cotton wool and ZIC-cotton HILIC both in ion-pairing (0.1% TFA) and non-ion pairing mode (0.5% FA), compared to a sample purified on a R2 reversed phase microcolumn (Figure 4). Different MALDI-TOF MS spectra patterns were observed for peptides enriched by cotton wool and ZIC-cotton both in ion-pairing and non-ion pairing mode, suggesting that different peptide profiles are obtained by using these types of HILIC stationary phases.

The peptides eluted from the cotton wool and ZIC-cotton microcolumns were subjected to digestion with PNGase A in the presence of $H_2{}^{18}O$ followed by trypsin-mediated back-exchange in $H_2{}^{16}O$ as described above and the deglycosylated peptides were subject to LC-MS/MS analysis on a Q-exactive orbitrap. We thus identified 78 different glycosylation sites assigned to 67 albumin glycoproteins (Table 1). In all identified glycosylation sites asparagine residues in NXS/T/C sequons were deamidated and labeled with ^{18}O. An MS/MS fragment ion spectrum of the deglycosylated peptide LDMDVTLV#NITR matched to TC246874 in the TIGR wheat database is shown as an example in Figure 5.

For both ion pairing and non ion pairing modes a higher number of glycopeptides were identified by ZIC-cotton as compared to cotton wool (Figure 6), presumably reflecting a broader glycopeptide selectivity of ZIC-cotton. A limited binding capacity of cotton wool was also considered, but the number of glycopeptides identified did not increase when approximately three fold larger amount of

cotton was used (data not shown). With both ZIC-cotton and cotton wool fewer glycopeptides were identified when using TFA as ion-pairing agent compared to non-ion pairing mode with FA (Figure 6 a, b). This may be related to a reduction in overall hydrophilicity due to neutralization of positively charged amino acids by TFA.

According to our survey of dbPTM[32] and uniprot databases, most glycosylation sites identified here have not before been experimentally verified and the data provided here thus provides novel insight into the composition of cereal proteomes that may be beneficial for further studies to gain deeper understanding on the structure and biological roles of individual proteins as discussed below.

Potential allergenicity of wheat glycoproteins

Major food allergens are glycoproteins[33] and most IgE-mediated allergic reactions involve water soluble proteins.[34] Generally an extensive immune cross-reactivity is found among different plant and insect species and the allergenicity of plant glycoproteins is related to the presence of xylose and α-1,3-linked core fucose residues which do not exist in mammalian glycoproteins.[35,36]

Recently, Palmisano et al. applied the Structural Database of Allergenic Proteins (SDAP http://fermi.utmb.edu/SDAP/) to screen identified grape glycoproteins for similarity to known allergens.[37] Here the same methodology revealed that 21 of the total 67 identified glycoproteins from wheat albumins have amino acid sequence similarity to food allergens (Table 1). Among these, dimeric α-amylase inhibitor (TC249880), 27K protein (TC250399) and serpin-N3.2 (TC265635) are similar to well known wheat allergens called Tri a TAI, Tri a 27.0101 and Tri a 33.0101 respectively.[38, 39] α-Amylase inhibitors and serpins have been identified as major allergens causing baker's asthma.[40] Thaumatin-like protein (U35_14143) and xylanase inhibitor proteins-1 (TC264929, TC264927) show similarity to allergens Mus a 4.0101 and Ziz m 1.0101, respectively. These proteins were recently identified as wheat allergens by immunoblotting with IgE from an atopic dermatitis patient.[41] Wheat β-D-glucan exohydrolase (TC250803), triticin (TC264477) and globulin-3A (TC234045) also reported to bind IgE[42] are homologs to Asp n 14, Pis v 2, and Ses i 3, respectively. Fructan 1-exohydrolase (TC253637), globulin-1S (TC246874), polygalacturonase-like protein (TC267315, U35_5393), and neutral ceramidase (U35_4630) show similarity to the

allergens Lyc e 2.0102, Cor a 11, Phl p 13, and Pha a 5, respectively. Since cross-reactivity among different plant species can be expected, these glycoproteins are obvious targets in the identification of novel wheat allergens.

Wheat β-glucosidase does not have similarity to any allergenic proteins in the SDAP database, but was reported to bind IgE.[41, 42] This underscores that identification of wheat allergens is still at a preliminary stage and allergenic proteins may not yet be available in the database. A 36 kDa allergen causing baker's asthma was identified as seed-specific peroxidase from wheat flour[43] and wheat peroxidase 1 is also identified as allergen.[41, 44] Six peroxidases were identified as glycoproteins in the present study, but did not match allergens in SDAP. However, a barley homologue (gi|326513532) to the identified putative peroxidase (U35_22623) has similarity to the Tri a gliadin allergen. These data provide a basis for further study of differences among allergenic and non-allergenic isoforms of cereal peroxidases.

The proteins identified here may have potential as natural immunomodulatory glycoproteins to enhance immune responses. One of the most useful approaches to modulate the immune system has been to identify natural compounds, such as phytoglycoproteins[45] due to their immunogenic carbohydrate epitopes. In order to evaluate the allergenic potential of the identified glycoprotein targets further analysis of possible IgE binding via carbohydrate moieties on glycoproteins can be performed for example by preincubation of the candidate allergenic glycoproteins with IgE from patient serum and testing for binding of polyclonal anti-horseradish peroxidase antibodies, which recognize *N*-linked glycan moieties with an α(1,3)-fucose branch.[46]

Glycoproteins related to production of beer

Many of the proteins present in beer are derived from cereal grains. Understanding the effects of these proteins on beer quality (e.g. foam and haze formation) requires information about characteristics and composition. Glycated and glycosylated proteins are reported to influence the quality of beer[47], but only a few studies are reported on cereal glycoproteins.[47] In addition, allergenicity in beer is also an issue of beer quality and glycoprotein identification may thus contribute for evaluation of beer quality.

Many proteins influencing beer quality have been identified by proteomic studies of the beer [48, 49], beer foam[50, 51, 52], and beer haze[53]. Most information about the influence of beer proteins has been reported for barley beers, while much less is available on wheat beer[48, 49], and to our knowledge no protein profiling was reported of wheat beer. Pathogenesis-related proteins (PRs) are considered to be the most important proteins for beer production due to their thermal stability and resistance toward proteolysis[54] and play roles in foam and haze formation[54]. This work showed that some of the PRs in wheat are glycosylated including dimeric amylase inhibitor (TC249880), serpin (TC265635), peroxidases (TC249600, TC251690, TC240657, U35_22623 and U35_14350) and thaumatin-like protein (U35_14143).

CONCLUSION

The combination of ZIC-HILIC and cotton wool provides a novel advancement in the use of HILIC-based stationary phases for enrichment of glycopeptides from complex protein mixtures. This is to our knowledge one of the first examples where HILIC together with enzymatic deglycosylation in $H_2{}^{18}O$ has been applied for large-scale glycopeptide identification in the area of plant glycoproteomics. Production of biopharmaceutical proteins in plants offers tremendous advantages in terms of cost-effectiveness[55]. However, engineering of the glycoprotein processing machinery is necessary to avoid potential problems with IgE cross-reactivity. To meet this challenge, futher development of these bioanalytical tools will be persued in order to gain further insight into the complexity of glycoproteomes.

ACKNOWLEDGEMENT

Birgit Andersen is acknowledged for technical assistance with mass spectrometric analysis and Louise Helsted and Julie Henriksen for providing the wheat albumin sample. We are grateful to Sequant (Umeå, Sweden) for providing ZIC-HILIC materials resins. Plaipol Dedvisitsakul is supported by a personal Ph.D. scholarship from the Thai Royal Government.

REFERENCES

1. Spiro, R. G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **2002**, *12*, 43R–56R.

2. Fitchette, A. C.; Dinh, O. T.; Faye, L.; Bardor, M. Plant proteomics and glycosylation. *Methods Mol. Biol.* **2007**, *355*, 317–342.

3. Bause, E.; Hettkamp, H. Primary structural requirements for *N*-glycosylation of peptides in rat liver. *FEBS Lett.* **1979**, *108*, 341−344.

4. Bause, E.; Legler, G. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr (Ser) for the *N*-glycosylation step during glycoprotein biosynthesis. *Biochem. J.* **1981**, *195*, 639−644.

5. Helenius, A.; Aebi, M. Roles of *N*-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **2004**, *73*, 1019–1049.

6. Song, W.; Henquet, M. G. L.; Mentink, R. A.; van Dijk, A. J.; Cordewener, J. H. G.; Bosch, D.; America, A. H. P.; van der Krol, A. R. *N*-glycoproteomics in plants: perspectives and challenges. *J. Proteomics* **2011**, *74*, 1463–1474.

7. Kimura, Y.; Hess, D.; Sturm, A. The *N*-glycans of jack bean alpha-mannosidase. Structure, topology and function. *Eur. J. Biochem*. **1999**, *264*, 168–175.

8. Lige, B.; Shengwu, M.; van Huystee, R. The effects of the site-directed removal of *N*-glycosylation from cationic peanut peroxidase on its function. *Arch. Biochem. Biophys*. **2001**, *386*, 17–24.

9. Lim, J. M.; Aoki, K.; Angel, P.; Garrison, D.; King, D.; Tiemeyer, M.; Bergmann, C.; Wells, L. Mapping glycans onto specific *N*-linked glycosylation sites of *Pyrus communis* PGIP redefines the interface for EPG-PGIP interactions. *J. Proteome Res.* **2009**, *8*, 673–680.

10. Geyer, H.; Geyer, R. Strategies for analysis of glycoprotein glycosylation. *Biochim. Biophys. Acta* **2006**, *1764*, 1853–1869.

11. Minic Z.; Jamet, E.; Négroni, L.; der Garabedian, P. A.; Zivy, M.; Jouanin, L. A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J. Exp. Bot.* **2007**, *58*, 2503–2512.

12. Catala, C.; Howe, K. J.; Hucko, S.; Rose, J. K.; Thannhauser, T. W. Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* **2011**, *11*, 1530–1544.

13. Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R. Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **2003**, *21*, 660–666.

14. Zhou, Y.; Aebersold, R.; Zhang, H. Isolation of *N*-linked glycopeptides from plasma. *Anal. Chem.* **2007**, *79*, 5826–5837.

15. Picariello, G.; Ferranti, P.; Mamone, G.; Roepstorff, P.; Addeo F. Identification of *N*-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass

spectrometry. *Proteomics* **2008**, *8*, 3833-3847.

16. Hagglund, P.; Bunkenborg, J.; Elortz, F.; Jensen, O. N.; Roepstorff, P. A new strategy for identification of *N*-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res*. **2004**, *3*, 556–566.

17. Hägglund, P.; Matthiesen, R.; Elortza, F.; Højrup, P.; Roepstorff, P.; Jensen, O.N.; Bunkenborg, J. An enzymatic deglycosylation scheme enabling identification of core fucosylated N-glycans and O-glycosylation site mapping of human plasma proteins. *J. Proteome Res.* **2007**, *6*, 3021-3031.

18. Wohlgemuth, J.; Karas, M.; Eichhorn, T.; Hendriks, R.; Andrecht, S. Quantitative site-specific analysis of protein glycosylation by LC-MS using different glycopeptide-enrichment strategies. *Anal. Biochem*. **2009**, *395*, 178–188.

19. Wada, Y.; Tajiri, M.; Yoshida, S. Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem*. **2004**, *76*, 6560–6565.

20. Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M.; Wuhrer, M. Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal. Chem*. **2011**, *83*, 2492–2499.

21. Wuhrer, M.; de Boer, A. R.; Deelder, A. M. Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom. Rev.* **2009**, *28*, 192–206.

22. Mysling, S.; Palmisano, G.; Højrup, P.; Thaysen-Andersen, M. Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics. *Anal. Chem.* **2010**, *82*, 5598–5609.

23. Omaetxebarria, M. J.; Hägglund, P.; Elortza, F.; Hooper, N. M.; Arizmendi, J. M.; Jensen, O. N. Isolation and characterization of glycosylphosphatidylinositol-anchored peptides by hydrophilic interaction chromatography and MALDI tandem mass spectrometry. *Anal. Chem.* **2006**, 15, *78*, 3335−3341.

24. Gobom, J.; Nordhoff, E.; Mirgorodskaya, E.; Ekman, R.; Roepstorff, P. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **1999**, *34*, 105-16.

25. Osborne, T. B. The proteins of the wheat kernel. In *The Carnegie Institution of Washington*, Carnegie Institute: Washington, DC, USA, 1907; Publ 84, pp. 1−119.

26. Wuhrer, M.; Hokke, C. H.; Deelder, A. M. Glycopeptide analysis by matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry reveals novel features of horseradish peroxidase glycosylation. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1741–1748.

27. Prenner, C.; Mach, L.; Glossl, J.; Marz, L. The antigenicity of the carbohydrate moiety of an insect glycoprotein, honey-bee (Apis mellifera) venom phospholipase A2. The role of alpha 1,3-fucosylation of the asparagine-bound *N*-acetylglucosamine. *Biochem. J.* **1992**, *284*, 377−380.

28. Kaji, H.; Yamauchi, Y.; Takahashi, N.; Isobe, T. Mass spectrometric identification of *N*-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. *Nat. Protoc.* **2006**, *6*, 3019–3027.

29. Küster, B.; Mann, M. 18O-labeling of N-glycosylation sites to improve the identification of gel-separated glycoproteins using peptide mass mapping and database searching. *Anal Chem*. **1999**, *71*, 1431−1440.

30. Hajkova, D.; Rao, K. C. S.; Miyagi, M. The pH dependency of the carboxyl oxygen exchange reaction catalyzed by lysyl endopeptidase and trypsin. *J. Proteome Res.* **2006**, *5*, 1667–1673.

31. Angel, P. M.; Lim, J. M.; Wells, L.; Bergmann, C.; Orlando, R. A potential pitfall in 18O-based *N*-linked glycosylation site mapping. *Rapid Commun. Mass Spectrom*. **2007**, *21*, 674–682.

32. Lee, T. Y.; Huang, H. D.; Hung, J. H.; Huang, H. Y.; Yang, Y. S.; Wang, T. H. dbPTM: an information repository of protein post-translational modification. *Nucleic Acids Res*. **2006**, *34*, D622–627.

33. Lack, G. Clinical practice: food allergy. *N. Engl. J. Med*. **2008**, *359*, 1252–1260.

34. O'Neil, C. E.; Lehrer, S. B. Occupational reactions to food allergens. In *The Food Allergy: Adverse Reactions to Foods and Food Additives*. 2nd ed.; Metcalfe, D. D., Sampson, H. A., Simon, R. A., Eds.; Blackwell Science: Cambridge, MA, USA, 1997; pp. 311−335.

35. Bardor, M.; Faveeuw, C.; Fitchette, A. C.; Gilbert, D.; Galas, L.; Trottein, F.; Faye, L.; Lerouge, P. Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose. *Glycobiology*. **2003**,13, 427−434.

36. Westphal, S.; Kolarich, D.; Foetisch, K.; Lauer, I.; Altmann, F.; Conti, A.; Crespo, J.F.; Miranda, E.E.; Vieths, S.; Scheurer, S. Molecular characterization and allergenic activity of Lyc e 2 (β-fructofuranosidase), a glycosylated allergen of tomato. *Eur. J. Biochem*. **2003**, *270*, 1327–1337.

37. Palmisano, G.; Antonacci, D.; Larsen, M. R. Glycoproteomic profile in wine: a "sweet" molecular renaissance. *J. Proteome Res.* **2010**, *9*, 6148–6159.

38. Sander, I.; Rozynek, P.; Rihs, H. P.; van Kampen, V.; Chew, F. T.; Lee, W. S.; Kotschy-Lang, N.; Merget, R.; Bruning, T.; Raulf-Heimsoth, M. Multiple wheat flour allergens and cross-reactive carbohydrate determinants bind IgE in baker's asthma. *Allergy* **2011,** 66, 1208-1215

39. Kimoto, M.; Suzuki, M.; Komiyama, N.; Kunimoto, A.; Yamashita, H.; Hiemori, M.; Takahashi, K.; Tsuji, H. Isolation and molecular cloning of a major wheat allergen, Tri a Bd 27K. *Biosci Biotechnol Biochem*. 2009, 73, 85−92.

40. Pastorello, E. A.; Farioli, L.; Conti, A.; Pravettoni, V.; Bonomi, S.; Iametti, S.; Fortunato, D.; · Scibilia, J.; Bindslev-Jensen, C.; Ballmer-Weber, B.; Robino, A. M.; Ortolani, C**.** Wheat IgE-mediated food allergy in European patients: α-amylase inhibitors, lipid transfer proteins and low-molecular-weight glutenins. Allergenic molecules recognized by double-blind, placebo-controlled food challenge. *Int. Arch. Allergy Immunol*. **2007**, *144*, 10–22.

41. Sotkovsky, P.; Sklenar, J.; Halada, P.; Cinova, J.; Setinova, I.; Kainarova, A.; Golias, J.; Pavlaskova, K.; Honzova, S.; Tuckova, L. A new approach to the isolation and characterization of wheat flour allergens. *Clin. Exp. Allergy* **2011**, *41*, 1031–1043.

42. Larré, C.; Lupi, R.; Gombaud, G.; Brossard, C.; Branlard, G.; Moneret-Vautrin, D. A.; Rogniaux, H.; Denery-Papini, S. Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction. *J. Proteomics* **2011**, *74*, 1279–1289.

43. Sanchez-Monge, R.; Garcia-Casado, G.; Lopez-Otin, C.; Armentia, A.; Salcedo, G. Wheat flour peroxidase is a prominent allergen associated with baker's asthma. *Clin. Exp. Allergy* **1997**, *27*, 1130–1137.

44. Matsuo, H.; Uemura, M.; Yorozuya, M.; Adachi, A.; Morita, E. Identification of IgE-reactive proteins in patients with wheat protein contact dermatitis. *Contact Derm.* **2010***, 63*, 23–30.

45. Lee, J.; Lim, K. T. SJSZ glycoprotein (38 kDa) modulates expression of IL-2, IL-12, and IFN-γ in cyclophosphamide-induced Balb/c. *Inflamm. Res*. **2012**, *61*, 1319−1328.

46. Jin, C.; Bencúrová, M.; Borth, N.; Ferko, B.; Jensen-Jarolim, E.; Altmann, F.; Hantusch, B. Immunoglobulin G specifically binding plant N-glycans with high affinity could be generated in rabbits but not in mice. *Glycobiology* **2006**, *16*, 349-357.

47. Laštovičková, M.; Bobálová, J. MS based proteomic approaches for analysis of barley malt. *J. Cereal Sci.* **2012**, *56*, 519–530.

48. Perrocheau, L.; Rogniaux, H.; Boivin, P.; Marion, D. Probing heat-stable water-soluble proteins from barley to malt and beer. *Proteomics* **2005**, *5*, 2849–2858.

49. Konečná, H.; Müller, L.; Dosoudilová, H.; Potěšil, D.; Buršíková, J.; Sedo, O.; Márová, I.; Zdráhal, Z. Exploration of beer proteome using OFFGEL prefractionation in combination with two-dimensional gel electrophoresis with narrow pH range gradients. *J. Agric. Food Chem.* **2012**, *60*, 2418–2426.

50. Iimure, T.; Takoi, K.; Kaneko, T.; Kihara, M.; Hayashi, K.; Ito, K.; Sato, K.; Takeda, K. Novel prediction method of beer foam stability using protein Z, barley dimeric alpha-amylase inhibitor-1 (BDAI-1) and yeast thioredoxin. *J. Agric. Food Chem*. **2008**, *56*, 8664–8671.

51. Hao, J.; Li, Q.; Dong, J. J.; Yu, J .J.; Gu, G. X.; Fan, W.; Chen, J. Identification of the major proteins in beer foam by mass spectrometry following sodium dodecyl sulfate- polyacrylamide gel electrophoresis. *J. Am. Soc. Brew. Chem.* **2006**, *64*, 166–174.

52. Okada, Y.; Iimure, T.; Takoi, K.; Kaneko, T.; Kihara, M.; Hayashi, K.; Ito, K.; Sato, K.; Takeda, K. The influence of barley malt protein modification on beer foam stability and their relationship to the barley dimeric alpha-amylase inhibitor-I (BDAI-I) as a possible foam-promoting protein. *J. Agric. Food Chem*. **2008**, *56*, 1458–1464.

53. Iimure, T.; Nankaku, N.; Watanabe-Sugimoto, M.; Hirota, N.; Tiansu, Z.; Kihara, M.; Hayashi, K.; Ito, K.; Sato, K. Identification of novel haze-active beer proteins by proteome analysis. *J. Cereal Sci*. **2009**, *49*, 141–147.

54. Stanislava, G. A review: the role of barley seed pathogenesis-related proteins (PRs) in beer production. *J. Inst. Brew*. **2010**, *116*, 111–124.

55. Hefferon, K. L. Recent Advances in Virus Expression Vector Strategies for Vaccine Production in Plants. *Virol Mycol.* **2012**, 1, 105.

Table 1: *N*-glycoproteins and their glycosylation sites identified in wheat albumin proteins

Accession number[a]	Identified protein homologue		Identified glycopeptiddes[d]			Method[e]	SDAP homolog allergen (score, e-value)[f]
	Protein name[b]	Accession number[c]	Peptide Score	Expect-value	Sequence		
TC250399	27K protein [*Triticum aestivum*]	gi|30793446	33	0.22	R.GH#NLSLEYGK.Q	zcI, zcN	Tri a 27.0101 (349, 7.4e-98)
TC250401	27K protein [*Triticum aestivum*]	gi|30793446	41	0.028	R.GH#NLSLEYGR.Q	zcI, zcN, cN	Tri a 27.0101 (363.6,3.0e-102)
			35	0.067	R.DGLLDAA#NLTLVPYGNAVVR.N	zcI, zcN	
TC270707	Aspartic proteinase nepenthesin-1 [*Triticum urartu*]	gi|473892189	37	0.047	K.SL#NFSGTLR.L	zcN	-
TC250803	Beta-D-glucan exohydrolase [*Triticum aestivum*]	gi|20259685	23	3	R.E#NATAEAMSK.Y	zcN, zcI	Asp n 14 (71.0, 3.3e-13)
TC266747	Beta-glucosidase [*Hordeum vulgare*]	gi|193073259	72	2.50E-05	R.LPGFST#NESR.M	zcI,cN	-
			30	0.31	R.YGNPTMILSENGMDQPG#NVSIADGVHDTVR.I	zcN	
TC270421	Bifunctional purple acid phosphatase 26 [*Aegilops tauschii*]	gi|475524769	48	0.081	R.MT#NYTFYDYR.S	zcN, cN, cI	-
TC262722	Calreticulin-like protein [*Triticum aestivum*]	gi|56606827	33	0.11	K.KDENMAGEW#NHTSGK.W	zcI, zcN	Pench 31.0101 (87.2, 1.7e-18)
TC266130	Chitinase 1 [*Triticum urartu*]	gi|473983657	61	0.00024	K.TG#NVTGLLSPDQGISGAK.E	cN, cI	-
TC249880	Dimeric alpha-amylase inhibitor [*Triticum aestivum*]	gi|108597921	57	0.00067	R.LQC#NGSQVPEAVLR.D	zcI, zcN, cN, cI	Tri a TAI (135.5, 3.2e-34)
U35_20390	Diphosphonucleotide phosphatase 1 [*Oryza sativa*]	Q10Q09	35	0.036	R.GSI#NTTYQLVK.D	zcN	-
TC269625	Early nodulin-like protein 2 [*Aegilops tauschii*]	gi|475568923	68	0.00016	R.FQI#NDTIVFAR.G	zcN	-
TC253637	Fructan 1-exohydrolase w2 [*Triticum aestivum*]	gi|75297789	58	0.0019	R.AHIYAFN#NGSATVR.V	zcN	Lyc e 2.0102 (360.7, 1.6e-100)
TC246874	Globulin-1 S allele [*Triticum urartu*]	gi|474323981	71	2.20E-05	K.LDMDVTLV#NITR.G	zcI, zcN, cN,cI	Cor a 11 (164.3, 1.3e-41)
TC234045	Globulin-3A [*Triticum aestivum*]	gi|390979705	34	0.38	R.VAVA#NITPGSMTAPYLNTQSFK.L	zcI, zcN, cN, cI	Ses i 3 (69.1, 8.6e-13)
U35_16369	High pI alpha-glucosidase [*Hordeum vulgare*}]	gi|8547062	30	0.17	R.YGYL#NVSDLER.V	zcN	-
			34	0.049	R.V#NFTAAELRPFVDR.L	ZcN	
TC271002	Hypothetical protein Os08g0114400 [*Oryza sativa*]	gi|115474505	22	5.1	R.FS#NVTDR.L	zcI	-
			41	0.075	R.AA#NATFPR.T	zcI	
TC263334	hypothetical protein F775_26385 [*Aegilops tauschii*]	gi|475602513	48	0.0036	R.LGSHASASQ#NVSVRPGSLYALTFAATR.T	zcN	-
TC266588	Lysosomal alpha-mannosidase [*Triticum urartu*]	gi|474367528	40	0.082	R.VNDFVAAALAQA#NVTR.T	zcI, zcN	-
			30	0.42	K.ELSTEIITSMAT#NK.T	zcI	-

Accession number[a]	Identified protein homologue		Identified glycopeptiddes[d]			Method[e]	SDAP homolog allergen (score, e-value)[f]
	Protein name[b]	Accession number[c]	Peptide Score	Expect-value	Sequence		
TC252180	Lysosomal Pro-X carboxypeptidase [*Triticum urartu*]	gi|326499536	27	0.3	K.#NISSSIIALVTEK.G	zcN	-
			60	0.00036	K.YLL#NDTFWR.R	zcN	
TC267299	Monocopper oxidase-like protein SKU5 [*Aegilops tauschii*]	gi|475598907	52	0.0022	R.FV#NESLWTR.V	zcN, cN, cI	-
TC238592	N-carbamyl-L-amino acid amidohydrolase-like protein[Oryza sativa]	gi|108862917	43	0.015	R.TVDNH#NVSFVDAADSAGYK.I	zcN	-
U35_4630	Neutral ceramidase [*Triticum aestivum*]	gi|161702907	47	0.011	R.KGSTV#NATFYSACPR.N	zcI	Pha a 5 (34.7, 7.3e-03)
TC238385	Nicastrin [*Triticum urartu*]	gi|473930622	44	0.0088	R.FYHSHLDSPANI#NSSSIAAAAALVAR.S	zcN	-
U35_14350	Peroxidase 1 [*Triticum urartu*]	gi|474074307	26	0.5	R.VGFY#NTTCPNAEAIVR.Q	zcN	-
			28	0.28	R.DSV#NITGSNAFYQVPSGR.R	zcN	-
TC249600	Peroxidase 12 [*Aegilops tauschii*]	gi|475575998	30	0.28	R.EGLFVSDQDLFT#NATTRPIVER.F	zcN	-
TC251690	Peroxidase (predicted protein)[*Hordeum vulgare*]	gi|326516688	59	0.00046	R.LAA#NCSAAPSR.L	zcI, zcN	-
TC240657	Peroxidase 7 [*Triticum monococcum*]	gi|57635159	56	0.00059	R.A#NASALIR.L	zcN	-
TC246864	Predicted phosphatidylinositol/phosphatidylglycerol transfer protein [*Triticum urartu*]	gi|474063325	41	0.07	K.ISATTDK#NITEGK.L	zcN	-
TC238335	Predicted protein (band 7 domain of flotillin like proteins) [*Hordeum vulgare*]	gi|326532692	40	0.027	R.FIESIA#NNTK.I	zcN	-
TC267933	Predicted protein (Aldose-1-epimerase-like)[*Hordeum vulgare*]	gi|357163924	44	0.046	R.M#NATALDK.A	zcI, zcN, cN	-
TC236517	Predicted protein (containing DOMON-like domain)[*Hordeum vulgare*]	gi|326499558	51	0.0025	R.TY#NITGYVPLGK.A	cN	-
CK161542	Predicted protein (LysM domain-containing protein) [*Hordeum vulgare*]	gi|326504418	63	0.00039	R.FAC#NATAPR.A	zcI, zcN	-
			20	3.4	R.ASTCQALVSYSPP#NATTLAAVR.A	zcN	
TC254902	Predicted protein (peroxidase) [*Hordeum vulgare*]	gi|326497837	50	0.0031	K.VFVQQYAD#NR.T	zcI, zcN	-
U35_22623	Predicted protein (Peroxidase)[*Hordeum vulgare*]	gi|326513532	45	0.0046	R.TDAVLVQ#NATTR.A	zcI	-
U35_18540	Predicted protein (probable inactive receptor kinase) [*Hordeum vulgare*]	gi|326506994	33	0.052	R.DF#NVSFNR.L	zcN	-
			32	0.074	R.L#NGSIPASLR.S	zcN	
TC263252	Predicted protein (Thioredoxin_like) [*Hordeum vulgare*]	gi|326491935	42	0.01	R.AGF#NDSQSDQATR.I	zcN	-
TC262818	Predicted protein (transmembrane 9 superfamily) [Hordeum vulgare]	gi|326513810	33	0.15	R.MHA#NESDVLLCR.S	zcI, zcN, cN	-
TC238262	Predicted protein [*Hordeum vulgare*]	gi|326493520	42	0.024	K.SL#NFSGTLR.L	zcN	-
TC262990	Predicted protein [*Hordeum vulgare*]	gi|326487362	46	0.0096	R.AY#NLSYVVGDAK.N	cI	-

Accession number[a]	Identified protein homologue		Identified glycopeptiddes[d]			Method[e]	SDAP homolog allergen (score, e-value)[f]
	Protein name[b]	Accession number[c]	Peptide Score	Expect-value	Sequence		
U35_2850	Predicted protein [*Hordeum vulgare*]	gi|326498187	31	0.074	K.A#NASVLNGGTRPLR.S	zcI	-
TC251613	Predicted protein [*Hordeum vulgare*]	gi|326515428	40	0.022	K.SELV#NGTVVK.G	cN	-
TC251898	Predicted protein [*Hordeum vulgare*]	gi|326533790	41	0.015	R.G#NLSFGALR.G	cN	-
BE423544	Predicted protein L-ascorbate oxidase [*Hordeum vulgare*]	dbj|BAJ99266.1	52	0.0064	K.FPGP#NITAR.A	zcI	-
TC236632	Predicted protein(Aldose epimerase) [*Hordeum vulgare*]	gi|326529733	66	0.00028	R.LVPNDG#NNTLHGGHR.G	zcI, zcN, cN	-
U35_1864	PREDICTED: aminoacylase-1-like [*Brachypodium distachyon*]	gi|357141869	33	0.063	R.#NLTFEFK.Q	zcI	-
TC253775	PREDICTED: probable polygalacturonase-like [Brachypodium distachyon]	gi|357114278	48	0.016	R.DVFAE#NVTMAGR.M	zcN	Cha o 2.0101 (62.4, 5.2e-11)
TC266673	Putative glycerophosphoryl diester phosphodiesterase 1 [*Aegilops tauschii*]	gi|475568570	42	0.019	K.VMIQST#NSSTLVK.F	zcN	-
TC269168	Putative inactive purple acid phosphatase 29 [*Triticum urartu*]	gi|473990469	43	0.03	K.#NYTNEEPR.Q	zcI, zcN	-
TC267315	Putative polygalacturonase [*Aegilops tauschii*]	gi|475320367	52	0.0078	R.EHTASLLDFGGVGDG#NTSNTAAFR.K	zcN	Phl p 13 (64.3, 1.4e-11)
U35_5393	Putative polygalacturonase [*Aegilops tauschii*]	gi|475571332	55	0.00057	K.#NVSNPIIIDQYR.S	cI	Phl p 13 (68.4, 5.7e-13)
U35_8881	Putative uncharacterized protein [*Hordeum vulgare*]	B9F6N7	38	0.017	R.#NGTYVIVSR.H	zcI	-
TC265635	Serpin-N3.2 [*Triticum aestivum*]	gi|379060943	33	0.19	K.DILPAGSID#NTTR.L	zcN	Tri a 33.0101 (564.6, 2.7e-162)
TC266500	Somatic embryogenesis receptor kinase 2 [*Aegilops tauschii*]	gi|475599204	58	0.0016	K.NI#NSSLADNPR.L	zcN	-
U35_1175	Sucrose:fructan 6-fructosyltransferase [*Triticum aestivum*]	gi|388330680	43	0.0099	R.SSD#NSSEMLHVLK.A	zcN	-
U35_14143	Thaumatin-like protein [*Triticum aestivum*]	gi|20257409	28	0.17	R.GQFEHNCPPT#NYSK.F	zcI	Mus a 4.0101 (205.9, 8.0e-55)
TC264477	Triticin [*Triticum aestivum*]	gi|171027826	43	0.014	R.LQSQNDDIIHV#NHTLK.F	zcI, zcN, cN, cI	Pis v 2.0101 (100.0, 2.5e-22)
TC264916	Vacuolar processing enzyme 2c [*Hordeum vulgare*]	gi|313660968	34	0.21	K.NFFAVLLG#NK.T	zcN, cN	-
U35_7090	Vacuolar-sorting receptor 1 [*Aegilops tauschii*]	gi|475599592	23	0.37	R.VEYLE#NITIPSALISK.S	zcN	-
TC247191	Vicilin-like antimicrobial peptides 2-2 [*Aegilops tauschii*]	gi|475617685	32	0.58	K.ILVTAL#NTTFDELER.I	zcI, zcN	Gly m Bd28K (202.3, 4.5e-53)
			34	0.35	K.HDYHPLDHSDIGVYLV#NLTAGSMMAPHVNPR.A	zcN	
TC247118	Vicilin-like antimicrobial peptides 2-2 [*Aegilops tauschii*]	gi|475617685	32	0.38	K.ILVTAL#NTTYDELAR.I	zcI, zcN, cN, cI	Gly m Bd28K (203.6, 1.4e-53)
			42	0.057	K.HDYEPLDHSDIGVYLV#NLTAGSMMAPHVNPR.A	zcN	
CD925597	Vicilin-like antimicrobial peptides 2-2 [*Triticum urartu*]	gi|473890163	53	0.0041	K.#NTTAVEQLLSAK.S	zcI, zcN, cN, cI	Gly m Bd28K (40.5, 7.1e-05)

Accession number[a]	Identified protein homologue		Identified glycopeptiddes[d]			Method[e]	SDAP homolog allergen (score, e-value)[f]
	Protein name[b]	Accession number[c]	Peptide Score	Expect-value	Sequence		
TC238565	Vicilin-like antimicrobial peptides 2-2 [*Triticum urartu*]	gi|473890163	57	0.01	R.GDVYNLEQGSILYIQSYP#NATR.E	zcI, zcN, cI	Len c1.0102 (60.4, 1.2e-10)
TC249265	Xylanase inhibitor [*Triticum aestivum*]	gi|226427708	49	0.0018	R.GFSCSNF#NFTR.A	zcN, cN	-
TC264929	Xylanase inhibitor protein I [*Triticum aestivum*]	gi|284178233	51	0.0022	K.QT#NYSSLIK.Y	zcI, zcN, cN	Ziz m 1.0101 (109.7, 2.2e-25)
TC264927	Xylanase Inhibitor Protein I [*Triticum aestivum*]	gi|73622088	57	0.00055	R.YSLPS#NR.S	zcN	Ziz m 1.0101 (97.9, 3.1e-22)
TC235814	Xylanase inhibitor TAXI-I [*Triticum aestivum*]	gi|23954367	58	0.00051	R.FVA#NTTDGSKPVSK.V	zcI, zcN, cN	-
TC236890	Xylanase inhibitor TAXI-IV [*Triticum aestivum*]	gi|56201272	47	0.0064	K.FVA#NTTDGNKPVSK.V	zcI, zcN, cN,cI	-

[a]Accession number of protein sequences obtained from the TIGR wheat or HarvEST35 databases.

[b]Protein names and [c]accession numbers of homologous sequences given by BLAST search against NCBI or UniProt databases.

[d]The identified glycopeptides with their mascot score and expect-value. The glycosylation sites are indicated by #. Annotated mass spectra are available in Supporting information Table S3.

[e]Methods for glycopeptide enrichment using ZIC-cotton in non ion-pairing mode (ZcN), ZIC-cotton in ion-pairing mode (ZcI), cotton wool in non ion-pairing mode (cN) and cotton wool in ion-pairing mode (cI).

[f]Based on homology search of TIGR wheat or HarvEST35 entries against SDAP database (https://fermi.utmb.edu/SDAP/).

FIGURE LEGEND

Figure 1. MALDI-TOF MS spectra of HRP-BSA tryptic peptides eluted from microcolumns packed with different resins. a) ZIC-cotton HILIC with mobile phase 80%ACN, 0.5%FA, b) cotton wool with mobile phase 80%ACN, 0.5%FA, c) ZIC-HILIC with mobile phase 80%ACN, 0.5%FA, d) ZIC-cotton HILIC with mobile phase 80%ACN, 0.1%TFA, e) cotton wool with mobile phase 80%ACN, 0.1%TFA, f) ZIC-HILIC with mobile phase 80%ACN, 0.1%TFA and, g) R2 reversed phase chromatography with mobile phase 0.1%TFA. The left panel shows the magnified spectra of peptides from m/z 950 to 2500. Ovals and stars indicate non-glycosylated peptides and glycopeptides, respectively. Theoretical masses of tryptic HRP and BSA peptides are listed in Supporting information Table S1 and S2, respectively. The mass spectra represent the average of > 400 laser-shots.

Figure 2. MS spectra of selected HRP-BSA tryptic peptides subjected to deglycosylation and $^{18}O/^{16}O$ back exchange. a) deglycosylation performed in ^{16}O water, b) deglycosylation performed in ^{18}O water, c) deglycosylation performed in ^{18}O water followed by $^{18}O/^{16}O$ C-terminal back exchange. The MS spectra were matched to the predicted glycopeptides SFAN#STQTFFNAFVEAMDR (m/z 2182.99), LYN#FSNTGLPDPTLN#TTYLQTLR (m/z 2642.35) and GLIQSDQELFSSPN#ATDTIPLVR (m/z 2501.29) and non-glycosylated DAIPENLPPLTADFAEDKDVCK (m/z 2458.18). # indicates glycosylation sites. The mass spectra represent the average of 1000 laser-shots.

Figure 3. Overview of the strategy developed to screen for *N*-glycosylation sites in wheat protein extracts.

Figure 4. MALDI-TOF MS spectra of tryptic peptides from wheat albumin proteins eluted from ZIC-cotton and cotton wool microcolumns. a) ZIC-cotton HILIC with mobile phase 80%ACN, 0.5%FA, b) cotton wool with mobile phase 80%ACN, 0.5%FA, c) ZIC-cotton HILIC with mobile phase 80%ACN, 0.1%TFA, d) cotton wool with mobile phase 80%ACN, 0.1%TFA, and e) R2 reversed phase chromatography with mobile phase 0.1%TFA. The mass spectra represent the average of > 600 laser-shots.

Figure 5. **Fragment ion spectrum of deglycosylated peptide LDMDVTLV#NITR matched to TC246874 in Tigr wheat database.** The # symbol indicates the *N*-linked asparagine converted to ^{18}O incorporated-aspartic acid by PNGase A.

Figure 6. Venn diagram summarizing the number of glycosylation sites identified after glycopeptide enrichment by ZIC-cotton and cotton wool in ion-pairing (mobile phase 80%ACN, 0.1%TFA) and non-ion pairing mode (mobile phase 80%ACN, 0.5%FA)

SUPPORTING INFORMATION

Supporting information Table 1: Predicted tryptic peptides of horseradish peroxidase.

Supporting information Table 2: Predicted tryptic peptides of bovine serum albumin.

Supporting information Table 3. MS/MS spectra of identified glycopeptides from wheat albumin showing the fragment ion series of the deglycosylated peptides labeled with ^{18}O.

Supporting information Figure 1. Representative images of custom-made HILIC microcolumns

Supporting information Figure 2. A comparison of three approaches to avoid ^{18}O-labeling at the C-terminal carboxyl group by using a protease inhibitor (Pelfabloc), heat inactivation and trypsin-catalyzed back exchange

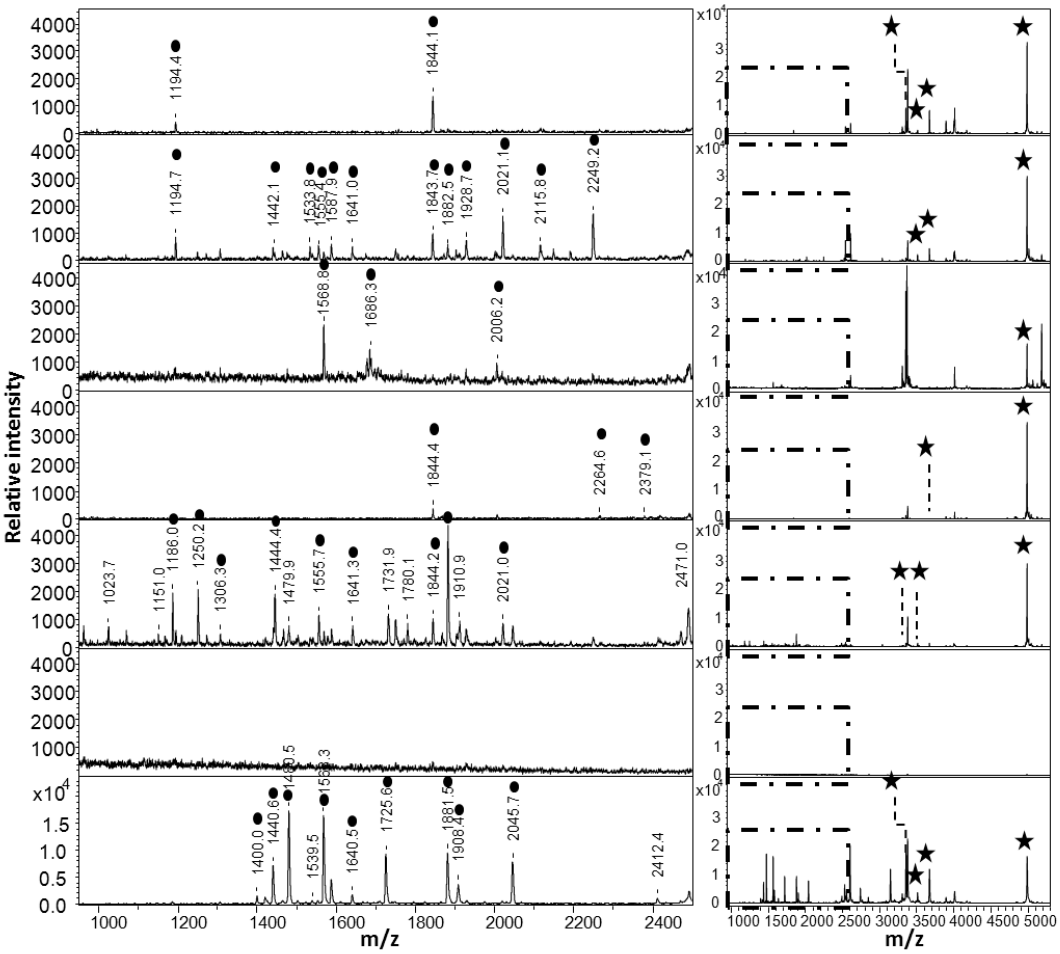


Figure 1. MALDI-TOF MS spectra of HRP-BSA tryptic peptides eluted from microcolumns packed with different resins. a) ZIC-cotton HILIC with mobile phase 80%ACN, 0.5%FA, b) cotton wool with mobile phase 80%ACN, 0.5%FA, c) ZIC-HILIC with mobile phase 80%ACN, 0.5%FA, d) ZIC-cotton HILIC with mobile phase 80%ACN, 0.1%TFA, e) cotton wool with mobile phase 80%ACN, 0.1%TFA, f) ZIC-HILIC with mobile phase 80%ACN, 0.1%TFA and, g) R2 reversed phase chromatography with mobile phase 0.1%TFA. The left panel shows the magnified spectra of peptides from m/z 950 to 2500. Ovals and stars indicate non-glycosylated peptides and glycopeptides, respectively. The mass spectra represent the average of > 400 laser-shots.

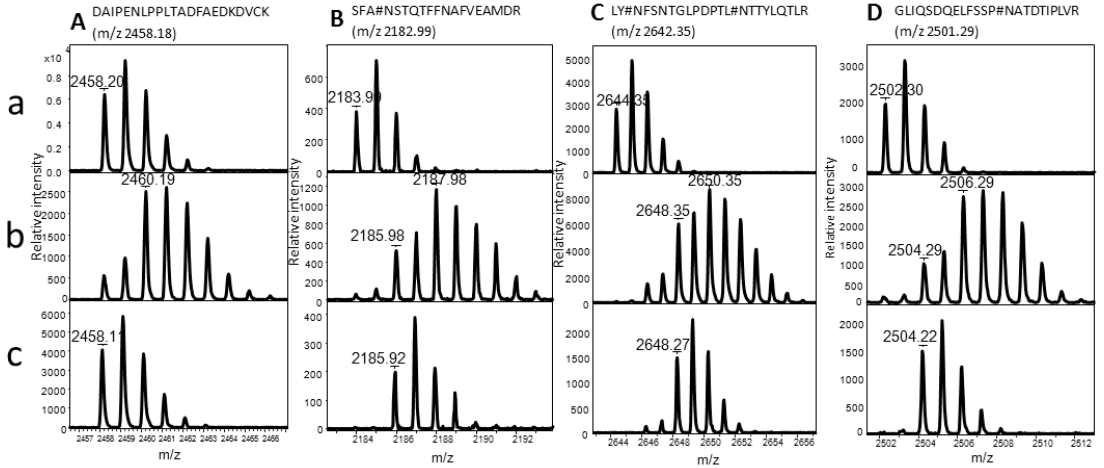


Figure 2. MS spectra of selected HRP-BSA tryptic peptides subjected to deglycosylation and $^{18}O/^{16}O$ back exchange. a) deglycosylation performed in ^{16}O water, b) deglycosylation performed in ^{18}O water, c) deglycosylation performed in ^{18}O water followed by $^{18}O/^{16}O$ C-terminal back exchange. The MS spectra were matched to the predicted glycopeptides SFAN#STQTFFNAFVEAMDR (m/z 2182.99), LYN#FSNTGLPDPTLN#TTYLQTLR (m/z 2642.35) and GLIQSDQELFSSPN#ATDTIPLVR (m/z 2501.29) and non-glycosylated DAIPENLPPLTADFAEDKDVCK (m/z 2458.18). # indicates glycosylation sites. The mass spectra represent the average of 1000 laser-shots.

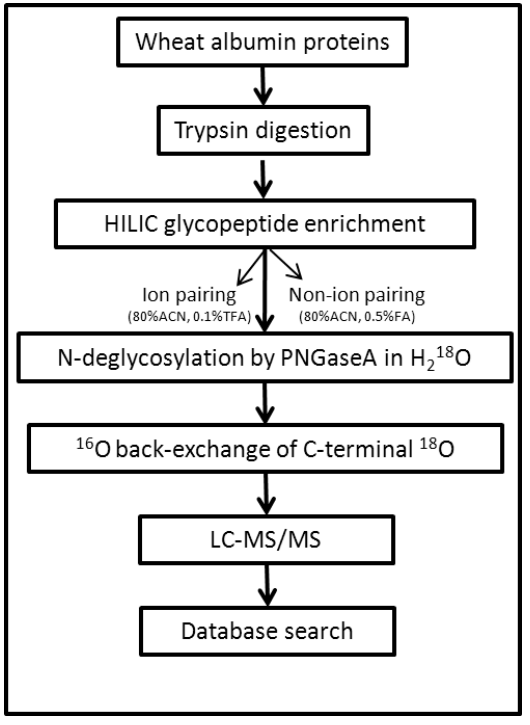


Figure 3. Overview of the strategy developed to screen for *N*-glycosylation sites in wheat protein extracts

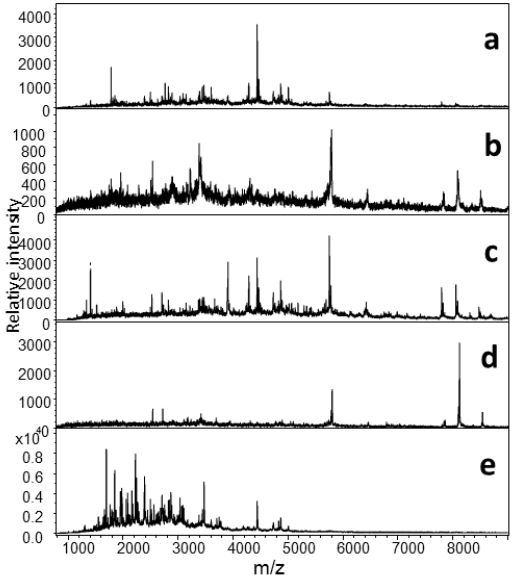


Figure 4. MALDI-TOF MS spectra of tryptic peptides from wheat albumin proteins eluted from ZIC-cotton and cotton wool microcolumns. a) ZIC-cotton HILIC with mobile phase 80%ACN, 0.5%FA, b) cotton wool with mobile phase 80%ACN, 0.5%FA, c) ZIC-cotton HILIC with mobile phase 80%ACN, 0.1%TFA, d) cotton wool with mobile phase 80%ACN, 0.1%TFA, and e) R2 reversed phase chromatography with mobile phase 0.1%TFA. The mass spectra represent the average of > 600 laser-shots.

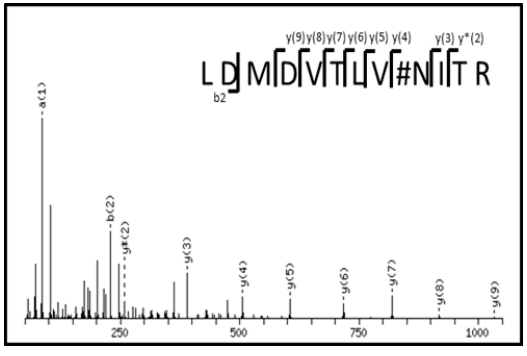


Figure 5. **Fragment ion spectrum of deglycosylated peptide LDMDVTLV#NITR matched to TC246874 in TIGR wheat database.** The # symbol indicates the *N*-linked asparagine converted to ^{18}O incorporated-aspartic acid by PNGase A. The * symbol indicates the ions that have lost ammonia.

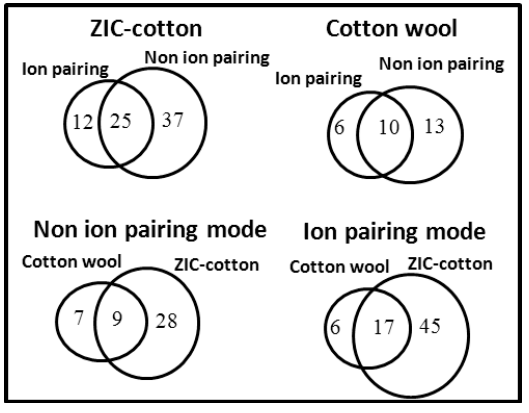


Figure 6. Venn diagram summarizing the number of glycosylation sites identified after glycopeptide enrichment by ZIC-cotton and cotton wool in ion-pairing (mobile phase 80%ACN, 0.1%TFA) and non-ion pairing mode (mobile phase 80%ACN, 0.5%FA)

Chapter 3 — Exploring Sub *N*-glycoproteome of Barley Aleurone Layer

Exploring Sub *N*-glycoproteome of Barley Aleurone Layer

Introduction

Protein *N*-glycosylation, a common post translational modification for proteins translocated through the secretory pathway, is known to affect protein folding, stability, activity, molecular interaction and functionality (Spiro et al., 2002). In addition, many glycoproteins have been shown to be involved in immunogenicity (Lack, 2008).

Due to the presence of highly abundant non-glycosylated proteins, detailed identification of glycoproteins present in complex biological sample is a challenge. Therefore, enrichment of glycoprotein/glycopeptides prior to mass spectrometry is required. Various strategies for glycoprotein/glycopeptide enrichment are based on lectin binding affinity (Zielinska et al., 2010), hydrazide chemistry (Zhang et al., 2003), size exclusion chromatography (Alvarez-Manilla et al., 2006) and hydrophilic interaction liquid chromatography (HILIC) (Calvano et al., 2008; Selby et al., 2008).

Currently very little information is available on glycosylated proteins from plants as compared to mammals. In plants, the first survey of the *N*-glycoproteome was performed on cell wall extracts from the stem of Arabidopsis by Minic et al., 2007 using affinity chromatography on concanavalin A-Sepharose followed by 2DE-separation (102 glycoproteins were identified). In another study by Catala et al., 2011, Con A was also applied for isolation of glycoproteins in ripe tomato fruit followed by trypsin digestion and glycopeptide captured with strong cation exchange chromatography (SCX). Zhang et al., 2011, recently used various lectins and boronic acid for glycoprotein enrichment from Arabidopsis cell wall protein extract an analysed by SDS-PAGE (127 glycoproteins were identified).

Barley (*Hordeum vulgare* L.), a member of the Triticeae grass tribe, is the major cereal plant of important in the brewery industry and used in several food products. The study of barley (i.e. proteins and glycosylations) could provide more insight into biological processes in barley which could be applied for future developments in brewing industry, human and animal nutritional aspects, plant breeding and/or cultivar identification.

Glycoprotein enrichment by ConA followed by SDS-PAGE, resulted in identification of some important glycoproteins such as serine carboxypeptidase and an α-amylase inhibitor

BMAI-I from barley grain as well as β-xylosidase from barley malt (Lastovickova et al., 2011).

Cereal aleurone layer proteins play a central role in seed germination as well as in industrial malting. During germination, gibberellic acid (GA) from the embryo triggers production of hydrolases in the aleurone layer that are released into the endosperm for degradation of storage proteins and polysaccharides to support seedling growth (Jacobsen and Verner, 1967, Schuurink R.C. et al., 1992, Finnie et al., 2011)

The in vitro culture approach of separated aleurone layer was developed by Chrispeels and Varner in 1967. It allows the study of germination, signaling, protein secretion and post-translational modification at a physiological and molecular level (Fath et al., 2002, Bethke et al., 1999, Christpeels et al., 1967, Bethke et al., 2002, Sticher et al., 1992). Several proteins produced from barley aleurone layer have been identified (Finnie et al., 2003, Finnie et al., 2011). However, the global barley aleurone layer glycoprotein profile has not yet been reported.

In the present study, HILIC-based chromatography using ZIC-cotton which is described in the previous chapter was applied to enrich glycopeptide from barley aleurone layer resulting in identification of 57 different glycosylation sites from 47 *N*-glycoproteins produced from barley aleurone layer. This is the first report providing a list of glycoproteins present in barley aleurone layer.

Material and Method

In vitro culture of barley aleurone layer and preparation of protein samples

The aleurone layers (ALs) of barley seeds (*Hordeum vulgare* L. cv. Himalaya) were isolated as described previously (Bønsager et al., 2007). For each treatment, 100 mg of ALs were incubated in 2 mL of a culture medium (20 mM sodium succinate, 20 mM $CaCl_2$, pH 4.2), and incubated at room temperature with gentle shaking for 24 h. After incubation, 5 µM gibberelic acid (GA) was added to the culture for samples referred to as GA-induced ALs. Two stress treatments were applied in ALs incubated in culture + GA 5 µM: either tunicamycin (added from a stock in methanol) or heat shock (40ºC, 4 h).

Aleurone layer cultures containing the extracellular proteins were harvested after 24h: ALs were washed four times with 2 mL of the baseline medium, frozen in liquid nitrogen, and stored at -80°C until use. Supernatants were pelleted by centrifugation to remove insoluble material, and frozen at -80º C until use. Intracellular proteins from ALs (100 mg) were extracted in 1 mL extraction buffer (5mM Tris*/*HCl, 1mM $CaCl_2$, pH7.5) with the protease inhibitor cocktail Complet (Roche, Germany) for 30 min in a cold chamber. Insoluble material was pelleted by centrifugation (10 min at 13000 rpm, 4ºC), and supernatants were stored at -80ºC until use (Østergaard et al., 2002). Protein concentration was determined using the Popov assay with bovine serum albumin as standard (Popov et al., 1975).

Identification of *N*-glycosylated proteins from barley aleurone layer complex samples and assignment of their glycosylation Sites

In-Solution Digestion and Hydrophilic Interaction Liquid Chromatography (HILIC) Protein samples from both intracellular and extracellular fractions were digested in-solution with trypsin as described previously (Hägglund et al 2004), and applied to a microcolumn prepared by a combination of ZIC-HILIC resin (Hägglund et al 2004) and cotton wool (Selman et al 2011). ZIC-cotton HILIC microcolumns were made by pushing a small piece of cotton wool (approximately 200 μg) taken from a cotton wool pad into a 20 μL GeLoader microtip (Eppendorf, Hamburg, Germany) and packing ZIC-HILIC resin (kind gift of Sequant, Umeå, Sweden; particle size 10 μm) on top of the cotton wool. The micro-column (Eppendorf, Hamburg, Germany) was equilibrated with 40 μL of 80% (v/v) acetonitrile (ACN), 0.5% (v/v) formic acid (FA). In-solution trypsin digested protein samples (20 μg in 85 μL) were dried (SPD 1010 Speed Vac, Thermo Scientific), redissolved in 40 μL 80% (v/v) ACN, 0.5% (v/v) FA and applied onto the equilibrated HILIC micro-columns. The column was subsequently washed twice with the same solution for equilibration followed by elution of bound peptides by 99.5% H_2O, 0.5% FA.

Endoglycosidase Digestion and LC-MS/MS analysis. Deglycosylation of the eluted peptides was performed in 100 mM ammonium acetate prepared in H_2O^{18} for 24 h at 37°C with 0.1 mU *N*-glycosidase A (PNGase A, Roache). To minimize ^{18}O- labeling at the C-terminal, the deglycosylated samples were dried and redissolved in 100 mM ammonium acetate prepared in normal water (^{16}O-water) and incubated for 24 h at 37°C in the presence of 0.2 μg of sequencing grade modified trypsin (Promega, Madison, WI). The peptides were cleaned up with StageTip C18-reversed phase column (Thermo Scientific) and analysed on EASY-Spray

column (Pepmap 3 µm, C18 15 cm x 75 µm) coupled on-line to a Q Exactive™ orbitrap (Thermo Scientific). The separation was performed using a flow rate of 300 nL min^{-1} and a gradient of solvents A (0.1% FA) and B (80% ACN, 0.1% FA) as follows: 5–40% B for 40 min; 40–100% B for 5 min; 100% B for 10 min. The LC-MS data were processed using Proteome Discoverer Software (version 1.3, ThermoFisher Scientific, CA, USA), then searched by MASCOT (Version 2.2.04, Matrix Science) against Barley TIGR Gene Index database (HvGI, version 11, http://compbio.dfci.harvard.edu/tgi/). Variable modifications were set to include deamidation of asparagine and glutamine with incorporation of ^{16}O (asparagine/glutamine + 0.98 Da) and ^{18}O (asparagine/glutamine + 2.99 Da) as well as one and two ^{18}O labels of C-terminal. Carbamidomethyl cysteine and methionine oxidation were set as fixed and variable modifications, respectively. Only glycopeptides containing the consensus tripeptide (NXS/T) in which the aspartic acid was tagged with ^{18}O atoms were accepted. Tentative consensus (TC) sequences obtained from the database were searched against NCBI (http://www.ncbi.nlm.nih.gov) by BLAST program to assign protein names.

Results and discussion

Detection of glycopeptides in complex biological samples by MS is challenging from a technical point of view because of ion-suppression effects from abundant non-glycosylated peptides in addition to the inherent high heterogeneity of glycopeptides due to variations of the glycan structure resulting in relatively low intensity of MS signals of the individual glycopeptides. In the present study a screening strategy is applied for identification of *N*-glycosylation sites in proteins from both intra and extracellular fraction of barley aleurone layer. To reduce complexity, non-glycosylated peptides are removed by HILIC-based glycopeptide enrichment after trypsin digestion using ZIC-cotton HILIC as described in the previous chapter. This approach coupled with site-specific glycosylation labeling and mass spectrometry allowed identification of 47 barley aleurone layer glycoproteins (Table 1). Identified barley aleurone layer glycoproteins and their glycosylation sites are presented and discussed in relation to these following aspects.

Biological processes associated with the identified barley aleurone layer glycoproteins

The glycoproteins identified in this study overlapped with proteins identified from the Con A binding fraction of soluble proteins extracted from tomato fruit pericarp (Catala et al., 2011)

and Arabidopsis stem cell wall (Minic et al., 2007). Carbohydrate active enzymes is the major group of identified glycoproteins in the present study including glycosyl hydrolases (GHs) belonging to families 3, 5, 27, 31, 32, 35, 38, 43, 51 and carbohydrate esterases (CE) from families 8. Proteases is the second largest group of identified glycoproteins including xylanase inhibitors, pepsin-like aspartic proteases, serine carboxypeptidases, *N*-acyl-L-amino-acid amidohydrolases, and subtilisin-like proteases. Proteins belonging to these GHs (except GH5) and CE8 families as well as many proteases identified in this study were also found in Con A binding fraction of soluble proteins extract from tomato fruit pericarp (Catala et al., 2011) and Arabidopsis cell wall (Minic et al., 2007). Serine carboxypeptidase and β-xylosidase (GH43) were also reported as Con A binding protein in barley grain and barley malt respectively (Lastovickova et al., 2011). This information supports the glycosylation of proteins belonging to these glycosyl hydrolase and protease groups. Other proteins for example peroxidase, glycine–aspartic acid–serine-leucine (GDSL) esterase/lipase, purple acid phosphatases were also found in the same study in tomato (Catala et al., 2011).

There are some glycoproteins identified in the present study that were not found in tomato pericarp (Catala et al., 2011) and Arabidopsis stem cell wall (Minic et al., 2007) for example embryo globulin and vicilin-like antimicrobial peptides 2-2 which are classified in the group of proteins having nutrient reservoir activities. These proteins may be therefore associated with specific cell types and biological pathways which are predominantly in grains.

N-linked glycosylations are found in proteins entering secretary pathway by targeting and co-translation into ER. To assess if these glycoproteins enters the secretary pathway, ER-targeting signal peptides predictions was also performed on full-length amino acid sequences of the identified glycoproteins using Signal P (www.cbs.dtu.dk/sevices/SignalP). Of 47 identified glycoproteins, 45 glycoproteins (95.7%) were predicted to contain a signal peptide. Only two glycoproteins (4.3%) did not contain a signal peptide e.g. DOMON domain-containing protein (gi|475558304) and an unknown protein (gi|326493228). These two proteins may be glycosylated via non-classical secretary pathway (Rose and Lee, 2010) as has been reported in animal (Nickel et al., 2009). However, an alternative pathway for secretion of glycosylation has not yet been reported in plants.

Potential glycoprotein allergens in barley aleurone layers

The same procedure for prediction of allergenic proteins as described in the previous chapter was also applied for prediction of potential glycoprotein allergens in barley aleurone layer using similarity searching against proteins sequences in structural database of allergenic proteins (SDAP). The amino acid sequences of the identified glycoproteins from barley alerone layers were searched against known allergens in SDAP. Many of the identified glycoproteins in the present study have similarity to known food allergens (29.8% of all glycoproteins identified in the present study; Table 1). The molecular functions of these potential allergenic proteins were also individually assigned and can be classified into 7 groups involved in glycoside hydrolases, proteases, nutrient reservoir, lipase, pectinesterase, calcium ion binding, and oxidation (Figure 1). Most of the potential allergenic glycoproteins in the present study have proteases (21.4%) and glycoside hydrolases (21.4%). Most of potential allergenic proteins found in this barley aleurone layer glycoprotein study have also been found in wheat albumins (chapter 2) for example β-D-glucan exohydrolase isoenzyme ExoI, embryo globulin, β-D-xylosidase, vicilin-like antimicrobial peptides 2-2 and calrecticulin. These results provide a set of protein targets for further studies on allergenic cross-reactivity found among Triticeae plant including wheat, barley and rye (Varjonen et al., 1994, Baldo et al., 1980). In addition, only few allergenic proteins have been identified and reported in barley. All of the potential allergenic proteins identified in the present study are not yet reported and can all be targets for identification of novel allergens in barley.

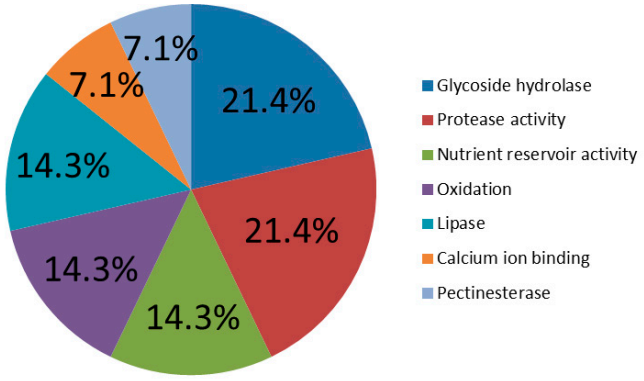


Figure 1. Molecular functions of the identified barley aleurone layer glycoproteins which are similar to food allergen in SDAP (Total 14 potential allergenic glycoproteins).

Potential targets for tunicamycin action

The endoplasmic reticulum (ER) is an essential organelle involved in several cellular functions including protein folding and secretion. The ER plays a role in cellular protein quality control by extracting and degrading proteins that are not correctly folded. This process, known as ER-associated degradation (ERAD), ensures that only properly folded and assembled proteins are transported to their final destinations (Tsai and Weissman 2010, Schroder and Kaufman, 2005).

Tunicamycin, an inhibitor of *N*-glycosylation process in ER, has been used to induce the accumulation of unfolded proteins in the lumen of ER and consequently to provoke an ER stress (Noh et al., 2003, Reis et al., 2011). If the unfolded proteins are persistent or excessive, ER stress triggers cell death, typically apoptosis (Xu et al., 2005). Links between ER stress and apoptosis have been reported in response to tunicamycin treatment in mammal cells, whereas in plants this correlation has been suggested, but the pathways of signal transduction remain unknown (Kamauchi et al., 2005, Reis et al., 2011).

Information of *N*-glycosylated proteins in barley is limited. Many *N*-linked glycoproteins and their glycosylation sites identified from both extracellular and intracellular fractions of barley aleurone layer in the present study confirms the presence of many *N*-glycosylated proteins which are potential targets for tunicamycin action in barley aleurone layer.

The identified glycoproteins in this present study included a part of further discussion in the manuscript in next chapter which involves effects of tunicamycin treatment and or heat shock on barley aleurone layer. In addition, identified glycoprotein from barley aleruone layer can serve as potential targets for tunicamycin action in further studies on the role of glycosylations.

Table 1. Identification of *N*-glycosylated proteins from barley aleurone layer and assignment of their glycosylation sites.

The identified glycopeptide sequences are obtained from LC-MS/MS spectra. The O^{18} –label asparagine indicating glycosylation sites are underlined in bold (**N**). Deamidation of asparagine and glutamine, carbamidomethyl cysteine, and methionine oxidation are assigned in lower case (n, q, c and m). Threoretical monoisotopic mass of peptides (MH+ [Da]) is indicated.The highest significant ion score among different protein samples and expected value of the identified peptides are also shown. Identifications in the different protein samples are indicated by (+). p.p.: predicted protein. Protein sequences containing the glycosylation sites and their accession numbers were obtained from BARLEY TIGR (version 11) or NCBI databases. The name of the identified glycoproteins and their GI accession number were given by BLAST search the protein sequences against NCBI database. Similar food allergens were given by homology search of the protein sequence against SDAP database (https://fermi.utmb.edu/SDAP/)

No	Protein identified	Accession	Signal peptide	Biological process	Allergen similarity in SDAP			Sequence	MH+ [Da]	Ion Score	Expected Value	Extra cellular	Intra cellular
					Allergens	Score	Expected value						
1	MtN19-like protein	gi|326513796/TC130798	yes	Stress response	-	-	-	VSDGEVLTIVS**N**YSSTR	1829.90	56	1.40E-03	+	-
								cDSV**N**VTIDEYGR	1594.66	79	1.90E-06	+	+
2	beta-D-Glucan Glucohydrolase Isoenzyme Exo1	gi|4566505/TC147448	yes	Carbohydrate metabolism (GH3)	Asp n 14	79.7	5.3e-16	HFVGDGGTVDGInE**N**NTIINR	2246.07	86	8.78E-06	+	-
								HFVGDGGTVDGINE**N**NTIINR	2245.08	91	2.94E-06	+	+
3	embryo globulin	gi|167004/TC139034	yes	Nutrient reservoir activity	Ses i 3	157.6	1.9e-39	VAVA**N**ITPGSMTAPYLNTQSFK	2313.18	81	2.15E-05	+	-
4	Endonuclease	gi|3242403/TC133379	yes	Nuclease activity	-	-	-	**N**ITDDWSSEEK	1326.56	75	1.86E-05	+	+
5	beta-D-xylosidase	gi|18025342/TC132616	yes	Carbohydrate metabolism (GH3)	Asp n 14	347.0	3.2e-96	GVQGYGMSGAI**N**SSDLEASACCK	2365.00	64	4.26E-04	+	-
6	p.p., thioredoxin like protein	gi|326491935/TC141383	yes	Oxidation	-	-	-	L**N**ASAVYPLLEK	1320.73	70	2.51E-05	+	+
								AGF**N**DSQSDQATR	1399.60	39	3.41E-02	+	-
7	p.p., cysteine-rich receptor protein kinase 10 (*Aegilops tauschii*)	gi|475529789/BU999544	yes	Phosphokinase activity	-	-	-	YSDQDFLA**N**VTDNR	1660.74	60	3.49E-04	+	-
8	p.p., protein disulphide isomerase	gi|326531866/TC146620	yes	Oxidation (involved in protein folding in ER)	Alt a 4	100.6	1.4e-22	LGPAVQ**N**LTAVDEAEK	1657.85	57	8.32E-04	+	-
								KLGPAVQ**N**LTAVDEAEK	1785.95	41	2.66E-02	+	-
9	cell wall invertase (*Triticum aestivum*)	gi|3342242/TC147887	yes	Carbohydrate metabolism (GH32)	Lyc e 2.0102	369.6	3.3e-103	SEDFL**N**WTK	1142.53	50	4.57E-03	+	+
								AHIYAFN**N**GSATVR	1523.75	78	3.82E-05	+	+
								AHIYAFn**N**GSATVR	1524.74	75	8.03E-05	+	+
10	oxalate oxidase	gi|2266668/TC148015	yes	Oxidation	Pru du 6.0201	41.7	4.7e-05	GGnTSTP**N**GSAVTR	1322.61	48	6.62E-03	+	+
11	p.p., xylanase inhibitor	gi|326487133/TC147599	yes	Protease acitivity (Aspatic protease)	-	-	-	GFScSNF**N**FTR	1339.57	68	3.03E-05	+	-
12	p.p., RNA binding protein	gi|326499558/TC148428	yes	protein-protein interactions	-	-	-	TY**N**ITGYVPLGK	1328.70	44	5.19E-02	+	-
13	p.p., secretory peroxidase	gi|326519386/BJ478399	yes	Oxidation	-	-	-	A**N**ASTLIR	848.47	43	1.69E-02	+	-
14	p.p., pepsin-like aspartic proteases	gi|326520291/TC149103	yes	Protease acitivity (Aspatic protease)	-	-	-	LPF**N**ASTFALK	1211.66	68	4.34E-05	+	+
15	p.p., vicilin-like antimicrobial peptides 2-2	gi|326513840/BQ757942	yes	Nutrient reservoir activity	Gly m Bd28K	108.8	5.5e-25	ILVTAL**N**TTYDELER	1753.92	71	2.97E-05	+	-
16	p.p., unannotated	gi|326493228/TC139840	no	Oxidation	-	-	-	GAFTNNQIWVTPY**N**K	1755.86	97	1.20E-07	+	-
								GAFTnNQIWVTPY**N**K	1756.85	53	2.85E-03	+	+
17	alpha-glucosidase	gi|3023275/TC132651	yes	Carbohydrate metabolism (GH31)	-	-	-	YGYL**N**VSDLER	1331.64	68	2.58E-04	+	+
18	p.p., purple acid phosphatase	gi|326514072/TC151097	yes	phosphatase	-	-	-	ISNVNY**N**ITSGNR	1454.71	67	1.20E-04	+	+
19	p.p., unannotated	gi|326533790/TC140254	yes	Unknown	-	-	-	G**N**LSFGALR	937.50	54	7.05E-03	+	-
								SYSt**N**LTFATVK	1347.67	61	1.38E-03	+	-
20	p.p., xylanase inhibitor	gi|326493694/TC147598	yes	Protease acitivity (Aspatic protease)	-	-	-	GLTcSNF**N**FTK	1291.59	68	6.19E-05	+	+
21	p.p., chitin elicitor-binding protein	gi|326511874/TC132570	yes	Peptidoglycan binding	-	-	-	VP**N**ATTYGALLAR	1349.74	55	1.17E-03	+	+
22	beta-galactosidase 15 (*Aegilops tauschii*)	gi|475522677/BF626888	yes	Carbohydrate metabolism (GH35)	-	-	-	GTDGPYFEPF**N**VSYDHR	2003.87	50	2.77E-03	+	-
23	pectinesterase (*Zea mays*)	gi|414879406/TC141846	yes	Carbohydrate metabolism (CE8)	Act d 7.0101	186.4	1.8e-48	V**N**GSDEYMAYYR	1470.61	94	7.05E-08	+	+
24	p.p., early nodulin-like protein 2	gi|326520537/TC147765	yes	copper binding protein	-	-	-	FQ**N**DTIVFVR	1354.73	44	1.16E-02	+	-

Continue

No	Protein identified	Accession	Signal peptide	Biological process	Allergen similarity in SDAP			Sequence	MH+ [Da]	Ion Score	Expected Value	Extra cellular	Intra cellular
					Allergen	Score	Expected value						
25	serine carboxypeptidase I	gi|2815493/TC132103	yes	Protease acitivity (Serine protease)	Api m 9.0101	117.8	1.4e-27	NLTSQGYR	941.46	42	2.54E-02	+	+
								NVSDYETGDLK	1243.56	69	4.52E-05	+	+
26	xylose isomerase	gi|6175480/TC146966	yes	Carbohydrate metabolism	-	-	-	IGFNGTLLIEPKPQEPTK	1985.09	48	2.80E-02	+	-
								IGFNGTLLIEPKPqEPTK	1986.08	40	1.88E-02	+	-
								KIGFNGTLLIEPKPQEPTK	2113.18	37	2.14E-02	+	-
27	arabinoxylan arabinofuranohydrolase isoenzyme AXAH-I	gi|13398412/TC134034	yes	Carbohydrate metabolism (GH51)	-	-	-	NTFDNTSR	957.42	42	1.66E-02	+	-
28	p.p., calrecticulin	gi|326534372/TC131540	yes	Calcium ions binding	Pen ch 31.0101	88.3	1.1e-18	DENMAGEWNHTSGK	1578.63	47	7.90E-03	+	-
29	p.p., alpha-mannosidase	gi|326502492/TC131793	yes	Carbohydrate metabolism (GH38)	-	-	-	ELSTEIITSMATNK	1540.77	68	7.72E-05	+	-
								EIESQLLPIANASSDLR	1858.97	50	3.42E-03	+	-
30	p.p., N-acyl-L-amino-acid amidohydrolase	gi|326520271/TC150266	yes	Protease activity	Phl p 3.0101	31.9	9.5e-03	NLTYQLMK	1013.53	40	3.30E-02	+	-
31	p.p., GDSL esterase/lipase	gi|326516774/TC150722	yes	Lipid metabolism	Hev b 4.0101	72.2	3.1e-14	AQTWcSPNATYcANR	1802.75	39	2.15E-02	+	+
32	p.p., purple acid phosphatase	gi|326499490/TC150391	yes	phosphatase	-	-	-	GSINTTYQLVK	1226.65	50	3.22E-03	+	-
33	p.p., unannotated	gi|326496609/TC142161	yes	unknown	-	-	-	GNcTVTTGLPGYR	1398.66	43	2.22E-02	+	-
34	p.p., purple acid phosphatase	gi|326500264/TC139780	yes	phosphatase	-	-	-	KGTFNQSMVADQMGK	1644.76	48	7.93E-03	+	-
35	p.p., GDSL esterase/lipase (*Brachypodium distachyon*)	gi|357129213/TC141158	yes	Lipid metabolism	Hev b 13	233.2	1.4e-62	YALIGNYSR	1059.53	41	3.35E-02	+	+
36	DOMON domain-containing protein (*A. tauschii*)	gi|475558304/TC141776	no	haem- and sugar-binding	-	-	-	SAAGIHSTSENLTYVR	1708.84	51	4.04E-03	+	-
37	p.p., serine carboxypeptidase S28	gi|326520173/TC132609	yes	Protease acitivity (Serine protease)	-	-	-	IEGDSSNcTSPEAVNTVR	1938.86	51	2.49E-03	+	+
38	glycerophosphoryl diester phosphodiesterase 1 (*Triticum urartu*)	gi|474317218/TC132548	yes	Lipidmetabolism	-	-	-	MDNcTTISDLFPK	1544.69	61	3.01E-04	+	-
39	p.p., alpha-galactosidase	gi|326508514/TC148946	yes	Carbohydrate metabolism (GH31)	-	-	-	NMTSQTIEILSNK	1481.74	48	7.82E-03	+	-
40	p.p., secretory peroxidase	gi|326499556/TC132105	yes	Oxidation	-	-	-	GcDASVLLNTTNSK	1482.70	59	5.85E-04	-	+
41	p.p., lipid transfer protein	gi|326489055/TC139446	yes	Lipid metabolism	-	-	-	AFNINVTDALR	1236.65	50	4.31E-03	-	+
42	p.p., transmembrane 9 family protein	gi|326513810/TC139887	yes	Unknown	-	-	-	MHANESDVLLcR	1447.66	40	4.39E-02	-	+
43	p.p., subtilisin-like protease-like	gi|326508452/TC140959	yes	Protease acitivity (Serine protease)	Cuc m 1	173.1	7.0e-44	CEGGSDFNSSACNR	1563.57	47	4.37E-04	-	+
44	p.p. N-acyl-L-amino-acid amidohydrolase	gi|326504618/TC147367	yes	Protease activity	-	-	-	NLTFEFK	901.45	40	3.40E-02	-	+
45	p.p., F775_26974 (*A. tauschii*)	gi|475459622/TC133182	yes	Unknown	-	-	-	LNTSATDLATR	1165.60	46	1.21E-02	-	+
46	1,4-beta-D-mannan endohydrolase precursor	gi|86169677	-	Carbohydrate metabolism (GH5)	-	-	-	VnNFTGVAYK	1116.55	47	2.72E-02	+	+
								VNNFTGVAYK	1115.56	78	2.02E-05	+	+
								nFNVSFNR	1001.46	48	1.35E-02	+	-
47	arabinoxylan arabinofuranohydrolase	gi|381142340/TC132139	yes	Carbohydrate metabolism (GH51)	-	-	-	GSANSTWGSVR	1124.53	53	2.26E-03	+	+

References

Alvarez-Manilla G.; Atwood J.; Guo Y.; Warren N. L.; Orlando R.; Pierce M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with ***N***-linked glycosylation sites. *J. Proteome Res.* **2006**, *5* 701–708.

Baldo, B. A.; Krilis, S.; Wrigley, C. W. Hypersensitivity to inhaled flour allergens. Comparison between cereals. *Allergy* **1980**, *35*, 45–56.

Bethke, P. C.; Lonsdale, J. E.; Fath, A.; Jones, R. L. Hormonally regulated programmed cell death in barley aleurone cells. *Plant Cell* **1999**, *11*, 1033–1045.

Bethke, P. C.; Fath, A.; Spiegel, Y. N.; Hwang, Y.; Jones, R. L. Abscisic acid, gibberellin and cell viability in cereal aluerone. *Euphytica* **2002**, *126*, 3–11.

Bønsager, B.C.; Finnie, C.; Roepstorff, P.; Svensson, B. Spatio-temporal changes in germination and radical elongation of barley seeds tracked by proteome analysis of dissected embryo, aleurone layer, and endosperm tissues. *Proteomics* **2007**, *7*, 4528–4540.

Calvano C. D.; Zambonin C. G.; Jensen O. N. Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry. *J. Proteomics* **2008**, *71* 304–317.

Catala C.; Howe K. J.; Hucko S.; Rose J. K.; Thannhauser, T. W. Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* **2011**, *11*, 1530–1544.

Chrispeels, M. J.; Varner, J. E. Gibberellic acid-enhanced synthesis and release of α-amylase and ribonuclease by isolated barley and aleurone layers. *Plant Physiol.* **1967**, *42*, 398–406.

Fath, A.; Bethkem, P. C.; Beligni, M. V.; Jones, R. L. Active oxygen and cell death in cereal aleurone cells. *J. Exp. Bot*. **2002**, *53*, 1273–1282.

Finnie, C.; Andersen, B.; Shahpiri, A.; Svensson, B. Proteomes of the barley aleurone layer: A model system for plant signalling and protein secretion. *Proteomics* **2011**, 11, 1595−1605.

Finnie, C.; Svensson, B. Feasibility study of a tissue-specific approach to barley proteome analysis: aleurone layer, endosperm, embryo and single seeds. *J. Cereal Sci.* 2003, 38, 217−227.

Hagglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; Roepstorff, P. A new strategy for identification of *N*-Glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res.* **2004**, *3*, 556–566.

Jacobsen, J. V.; Varner, J. E. Gibberellic acid-induced synthesis of protease by isolated aleurone layers of barley. *Plant Physiol.* **1967**, *42*, 1596–1600

Kamauchi S, Nakatani H, Nakano C, Urade R. Gene expression in response to endoplasmic reticulum stress in Arabidopsis thaliana. *FEBS J.* **2005**, 272, 3461–3476.

Lack, G. Clinical practice: food allergy. *N. Engl. J. Med*. **2008**, *359*, 1252–1260.

Lastovickova, M.; Smetalova, D.; Bobalova, J. The combination of lectin affinity chromatography, gel electrophoresis and mass spectrometry in the study of plant glycoproteome: Preliminary insights. *Chromatographia* **2011**, *73*, 113–122.

Minic Z.; Jamet, E.; Négroni, L.; der Garabedian, P. A.; Zivy, M.; Jouanin, L. A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J. Exp. Bot.* **2007**, *58*, 2503–2512.

Nickel, W.; Rabouille, C.; Mechanisms of regulated unconventional protein secretion. *Nat. Rev. Mol. Cell Biol.* **2009***, 10,* 148–155.

Noh, S. J.; Kwon, C.S.; Oh D.H.; Moon, J. S.; Chung, W. I. Expression of an evolutionarily distinct novel BiP gene during the unfolded protein response in Arabidopsis thaliana. *Gene* **2003**, 311, 81–91.

Østergaard, O.; Melchior, S.; Roepstorff, P.; Svensson B. Initial proteome analysis of mature barley seeds and malt. *Proteomics* **2002**, *2*, 733–799.

Popov, N.; Schmitt, M.; Schulzeck, S. M.; Matthies, H. Reliable micro method for determination of the protein content in tissue homogenates. *Acta Biol. Med. Ger.* **1975**, *34*, 1441–1446

Reis, P.; Rosado, G.; Silva, L.; Oliveira, L.; Oliveira, L.; Costa, M.; Alvim, F.; Fontes, E. The binding protein BiP attenuates stress-induced cell death in soybean via modulation of the N-rich protein-mediated signaling pathway. Plant Physiol **2011**, *157*, 1853-1865

Rose, J. K. C.; Lee, S.J.; Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome. *Plant Physiol.* **2010**, *153*, 433–436.

Schuurink, R. C.; Sedee, N. J. A.; Wang, M. Dormancy of the barley grain is correlated with gibberellic acid responsiveness of the isolated aleurone Layer *Plant Physiol.* **1992**, *100*, 1834–1839

Selby D. S.; Larsen M. R.; Calvano C. D.; Jensen O. N. dentification and characterization of ***N***-glycosylated proteins using proteomics. *Methods Mol. Biol.* **2008**, 484 263–276.

Selman, M. H. J.; Hemayatkar, M.; Deelder, A. M.; Wuhrer, M. Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal. Chem.* **2011**, *83*, 2492–2499.

Schroder, M.; Kaufman, R. J. ER stress and the unfolded protein response. *Mutat. Res.* **2005**, *569*, 29-63.

Spiro, R. G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **2002**, *12*, 43R–56R.

Sticher, L.; Jones, R. L. alpha-Amylase Isoforms are Posttranslationally Modified in the Endomembrane System of the Barley Aleurone Layer. *Plant Physiol*. **1992**, *98*, 1080−1086.

Tsai, Y. C.; Weissman, A. M. The Unfolded Protein Response, Degradation from the Endoplasmic Reticulum, and Cancer. *Genes & Cancer* **2010**, *1*, 764–778

Xu, C.; Bailly-Maitre, B.; Reed J. C. Endoplasmic reticulum stress: cell life and death decisions. *J. Clin. Invest.* **2005**, 115, 2656–2664

Varjonen, E.; Savolainen, J.; Mattila, L.; Kalimo, K. IgE-binding components of wheat, rye, barley and oats recognized by immunoblotting analysis with sera from adult atopic dermatitis patients. *Clin. Exp. Allergy* **1994**, *24*, 481–489.

Zielinska, D. F.; Gnad F.; Wiśniewski J. R.; Mann M. Precision mapping of an *in vivo* *N*-glycoproteome reveals rigid topological and sequence constraints. *Cell* **2010**, *141* 897–907.

Zhang, H.; Li X, J.; Martin, D. B.; Aebersold, R. Identification and quantification of ***N***-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **2003**, *21*, 660–666.

Zhang, Y.; Giboulot, A.; Zivy, M.; Valot, B.; Jamet, E.; Albenne, C. Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* **2011**, *72*, 1109–1123.

Chapter 4 — GA_3-induced aleurone layers responding to heat shock or tunicamycin provide insight into the *N*-glycoproteome, protein secretion and ER stress

Running head: Protein secretion and ER stress in aleurone layers

Corresponding author: Christine Finnie, Agricultural and Environmental Proteomics, Department of Systems Biology, Søltofts Plads, Building. 224, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. Tel +45 45252739, csf@bio.dtu.dk

Research area: Biochemistry and Metabolism

Title: GA_3-induced aleurone layers responding to heat shock or tunicamycin provide insight into the *N*-glycoproteome, protein secretion and ER stress

Authors: Gregorio Barba-Espín[1], Plaipol Dedvisitsakul[2], Per Hägglund[2], Birte Svensson[2], Christine Finnie[1].

Adresses: [1]Agricultural and Environmental Proteomics, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

[2]Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

One sentence summary: Overlapping responses of the aleurone layer proteome to tunicamycin and heat shock identify components of the plant protein secretory machinery

Footnotes:

This work was supported by the Carlsberg Foundation (to BS), the Danish Council for Independent Research | Natural Sciences and the Centre for Advanced Food Studies. The authors also thank the Thai Royal Government for a PhD stipend (to PD).

Corresponding author: Christine Finnie csf@bio.dtu.dk

ABSTRACT

The growing relevance of plants for production of recombinant proteins makes understanding the secretory machinery, including identification of glycosylation sites in secreted proteins, an important goal of plant proteomics. Barley (*Hordeum vulgare*) aleurone layers maintained *in vitro* respond to GA_3 by secreting an array of proteins and provide a unique system for analysis of plant protein secretion. Perturbation of protein secretion in GA_3-induced aleurone layers by two independent mechanisms: heat shock and tunicamycin treatment, demonstrated overlapping effects on both the intracellular and the secreted proteomes. Proteins in a total of 22 and 178 2D-gel spots changing in intensity in extracellular and intracellular fractions, respectively, were identified by mass spectrometry. Among these are proteins with key roles in protein processing and secretion, for example calreticulin, protein disulphide isomerase, proteasome subunits and isopentenyl diphosphate isomerase. Sixteen heat shock proteins (HSP) in 29 spots showed diverse responses to the treatments, with only a minority increasing in response to heat shock. The majority, all of which were small HSPs, decreased in heat shocked aleurone layers. Additionally, glycopeptide enrichment and *N*-glycosylation analysis identified 73 glycosylation sites in 65 aleurone layer proteins; 53 of the glycoproteins found in extracellular and 36 in intracellular fractions, respectively. This represents major progress in characterisation of the barley *N*-glycoproteome since only four of these sites were previously described. Overall, these findings considerably advance knowledge of the plant protein secretion system in general and emphasise the versatility of the aleurone layer as a model system for studying plant protein secretion.

INTRODUCTION

Plant proteins that are secreted to the apoplast have important functions in signalling, defence and cell regulation. The classical protein secretory pathway is less characterized in plants than in mammals or yeast, but is of growing interest due to the potential of plant systems as hosts for production of recombinant proteins (Erlendsson et al., 2010; De Wilde et al., 2013). Plant secretomics is therefore a rapidly expanding area applied to gain further insight into these processes (Agrawal et al., 2010; Alexandersson et al., 2013). Many secretory proteins contain putative *N*-glycosylation sites and the identification and characterisation of these sites is an important element in secretomics analysis. However, to date only a few plant glycoproteomes have been described (Fitchette et al., 2007; Minic et al., 2007; Palmisano et al., 2010; Melo-Braga et al., 2012; Zhang et al., 2012; Thannhauser et al., 2013).

The cereal aleurone layer is of major importance due to its central role in grain germination. Previous proteomic studies have been reported for aleurone layers dissected from mature (Finnie and Svensson, 2003) or germinating (Bønsager et al., 2007) barley or developing (Tasleem-Tahir et al., 2011) and mature (Laubin et al., 2008; Jerkovic et al., 2010; Meziani et al., 2012) wheat grains. The *in vitro* culture of isolated aleurone layers was developed by Chrispeels and Varner in 1967. Since then, this system has become an excellent tool for the study of germination signalling in response to phytohormones (Bush et al., 1986; Jones and Jacobsen, 1991; Bethke et al., 1997; Ishibashi et al., 2012). More recently it has been adopted as a unique system for analysis of plant secretory proteins (Hägglund et al., 2010; Finnie et al., 2011). Addition of GA_3 to the isolated aleurone layers induces synthesis and secretion of hydrolytic enzymes. In the *in vitro* system, these accumulate in the incubation buffer facilitating their identification and characterization using proteomics techniques. Thus, numerous secreted proteins with roles in hydrolysis of starch, cell wall polysaccharides and proteins could be identified (Finnie et al., 2011). Furthermore, several of the proteins were also detected in intracellular extracts from the same aleurone layers, presumably prior to their release. Many of the proteins appeared in multiple forms on 2D gels, and often with higher relative molecular mass (M_r) than expected suggesting the presence of posttranslational modifications (PTMs) (Finnie et al., 2011). Bak-Jensen et al (2007) and Finnie et al. (2011) observed a highly complex pattern of α-amylase-containing

spots in 2D-gels, which originated from only two AMY2 and two AMY1 gene products from a total of ten genes, most probably reflecting multiple forms due to PTMs.

In eukaryotic cells, proteins synthesised in the endoplasmic reticulum (ER) must be correctly folded and assembled before continuing in the secretory pathway. Pertubations of redox state, calcium regulation, glucose deprivation, and viral infection can lead to ER stress, triggered by accumulation of unfolded and misfolded proteins in the ER lumen. This provokes a triple response from the cell, consisting of an upregulation of chaperones and vesicle trafficking; down-regulation of genes encoding secretory proteins; and an up-regulation of proteins involved in ER-associated protein degradation (ERAD). In plants, the molecular mechanisms underlying ER stress in plants have yet to be fully resolved (Martínez and Chrispeels, 2003; Nagashima et al., 2011; Moreno et al., 2012).

Tunicamycin, an inhibitor of GlcNAc phosphotransferase, which catalyzes the first step in glycoprotein synthesis, has been used to induce ER stress by causing accumulation of unfolded proteins in the ER lumen (Noh et al., 2003; Kamauchi et al., 2005; Reis et al., 2011). If unfolded proteins are not removed, the prolonged stress will induce programmed cell death. Links between ER stress and apoptosis have been reported in response to tunicamycin treatment in mammalian cells, whereas in plants this correlation has been suggested, but the pathways of signal transduction remain unknown (Kamauchi et al., 2005; Reis et al., 2011).

In all plant tissues, heat shock induces the synthesis of a variety of heat shock proteins (HSPs), which are responsible for protein refolding under stress conditions (Craig et al., 1994) and translocation and degradation in a broad array of normal cellular processes (Bond and Schlesinger, 1986; Spiess et al., 1999). In barley aleurone layer heat shock selectively suppresses synthesis of secretory proteins including α-amylase, due to the selective destabilization of secretory protein mRNA (Belanger et al., 1986; Brodl and Ho, 1991). However, an acclimation effect has been described in aleurone cells after prolonged incubation at warm temperatures, resulting in a resumption of the protein secretory machinery (Shaw et al., 2003). The connection between heat stress response and ER stress has been well established in mammals and yeast, but scarce information is available in plants (Denecke et al., 1995).

Over the last years, the dual role of reactive oxygen species (ROS) has been established in plants: i.e. at higher concentrations, ROS act as toxic molecules damaging cellular

macromolecules and eventually causing cell death; but, at lower concentrations ROS seem to be necessary for seed germination and seedling growth by controlling the cellular redox status, regulating growth and protecting against pathogens (Bailly, 2004; Bailly et al., 2008). In the barley aleurone layer, GA_3 perceived at the plasma membrane induces ROS generation as a by-product from intense lipid metabolism, and the redox regulation of GA_3-induced response has been proposed (Maya-Ampudia and Bernal-Lugo, 2006). This suggests that during the secretory function of the tissue, moderate levels of ROS may be acting as cellular messengers. In aleurone cells, ROS, especially hydrogen peroxide (H_2O_2), are involved in the process of programmed cell death, but the molecular mechanisms remain unclear (Bethke and Jones, 2001; Ishibashi et al., 2012).

Until now none of the protein components of the ER stress pathways have been identified in barley, also little is known about glycosylation of barley proteins. In the present work, numerous *N*-glycoslation sites are identified and the effect of perturbing *N*-glycosylation and the secretory pathway by tunicamycin and heat shock treatments is analysed in GA_3-induced barley aleurone layers.

RESULTS

The Effect of Tunicamycin and Heat Shock on α-Amylase Production in GA_3-Induced Aleurone Layers

Since α-amylase is the most prominent secreted enzyme produced by aleurone cells in response to GA_3, a perturbation of the secretory machinery by tunicamycin or heat shock may affect the amount of this enzyme. Western blotting was used to detect α-amylase in the extracellular and intracellular fractions from aleurone layers incubated for 24 h. Untreated, GA_3-induced, and GA_3-induced aleurone layers in combination with heat-shock (HS) or tunicamycin (TN)-treatment were compared (Fig. 1). A decrease in α-amylase could be observed, both in the extracellular (Fig. 1A) and intracellular (Fig. 1B) fractions, when GA_3-induced cells were treated with 5 or 20 µg mL^{-1} TN. Higher tunicamycin concentrations did not elicit a further decrease in the amount of α-amylase (data not shown).

For HS treatment, a temperature of 40ºC was applied for the final 4 h of the 24 h incubation period. The HS treatment caused a decrease in α-amylase in the intracellular

fraction (Fig. 1B) , and in α-amylase secreted during the final 4 h, when compared to GA_3-induced aleurone cells (Fig. 1A). Longer incubation periods at 40ºC had no apparent effect on the amount of α-amylase (data not shown), most likely due to the acclimation response previously reported (Shaw et al., 2003).

GA_3-induced Cell Death Correlates with Enhanced Hydrogen Peroxide and Lipid Peroxidation

Both H_2O_2 content and thiobarbituric acid-reactive substances (TBARS) indicating lipid peroxidation were increased in extracts from GA_3-induced aleurone layers in comparison with untreated samples (Fig 2A). Treatment of GA_3-induced aleurone layers with 5 μg mL^{-1} TN or with HS did not result in further increases, but 20 μg mL^{-1} TN significantly enhanced the level of lipid peroxidation.

In order to examine whether TN and HS could affect viability of barley cells, cell death was monitored and quantified in intact aleurone layers by simultaneous staining of live and dead cells with fluorescent probes (Fig. 2B). The cell death quantitation reflected a correlation with the endogenous H_2O_2 and lipid peroxidation measurements. GA_3 treatment of barley aleurone cells resulted in increased cell death, while untreated aleurones maintained higher cell viability. Hence, after 24 h incubation, 45% of GA_3-induced cells were dead, whereas 81% of untreated cells remained alive (Fig. 2B). GA_3 + HS and GA_3 + 5 μg mL^{-1} TN did not result in increased cell death, whereas cell death was enhanced in GA_3 + 20 μg mL^{-1} TN samples (Fig. 2B). Treatments of GA_3-induced aleurone layers with 5 μg mL^{-1} TN and 4 h of HS were selected for comparative proteome analysis, since these affected protein secretion without provoking massive oxidative damage and cell death.

Fluorescent Glycoprotein Staining and Protein Identification

To visualise glycoproteins and assess the effect of tunicamycin treatment on protein glycosylation, extracellular proteins from GA_3-induced and GA_3 + 5 μg mL^{-1} TN-treated aleurone layers were resolved by 2D SDS-PAGE, and the glycoproteins were detected by Pro-Q® Emerald staining. As expected, TN greatly reduced glycosylation of the extracellular proteins (Fig. 3A). The fact that some highly abundant proteins did not react with the Pro-Q® Emerald stain whereas some low-abundance proteins

strongly reacted supports the specificity of Pro-Q® Emerald staining for glycoproteins. To reduce the risk of false positives, of the 53 fluorescent spots visible in all three GA_3 treated replicates, only the 23 intense spots missing after TN treatment (numbered spots, Fig. 3B) were analysed by mass spectrometry. Proteins were identified in 14 spots (Supplemental Table S1). Sequence analysis (http://www.cbs.dtu.dk/services/NetNGlyc/) confirmed the presence of at least one putative *N*-glycosylation site in each identified protein (data not shown).

Proteome Analysis of the Extracellular Fraction From *in vitro* Incubated Aleurone Layers.

Two-dimensional gel electrophoresis (2-DE; pH 3-10) was used to profile the extracellular protein fraction from the four aleurone layer treatments. Clear differences were apparent between 2-DE patterns of untreated and GA_3-induced aleurone layers, as well as between GA_3-induced aleurone layers and each stress treatment (Fig. 4A). In total, the average volumes of 35 spots varied by at least 1.5-fold among the four treatments. These spots were chosen for further analysis, and 22 were identified by mass spectrometry (Fig. 4B; Supplemental Table S2). As expected, most of them were α-amylases and proteases (Fig. 4C). The spots formed three clusters according to their appearance profiles (Fig. 5). Protein spots in Cluster A (13 spots) increased in response to GA_3, 11 contained α-amylases (AMY1 and AMY2) and one a cysteine proteinase (EP-B1). Moreover 4 spots decreased in GA_3 + TN samples (Cluster B). Here, one AMY2 spot migrated more slowly than the other identified AMY1 and AMY2 spots (Supplemental Table S2), indicating that this upper spot may be glycosylated. Finally five spots decreased in intensity in both treatments (Cluster C). Cluster C contains 4 protease spots (2 cysteine proteinase and 2 cathepsin) with higher experimental M_r than cysteine proteinase spots in Cluster A and B, again suggesting the presence of PTMs (Supplemental Table S2).

Proteome Analysis of Aleurone Layer Intracellular Extracts.

Intracellular water-soluble proteins from the four samples described above were also analysed using two-dimensional gel electrophoresis (Fig. 6A) and mass spectrometry. The image analysis showed 250 protein spots varying among the four treatments (Fig.

6B). Of these, 178 were identified (Fig 6B; Supplemental Table S3), and assigned to seven functional categories (Fig. 6C). Among the identified proteins 55% (98 protein spots) are involved in primary metabolism; while the remaining proteins were functionally assigned as chaperones (29), polysaccharide hydrolases (11), detoxification enzymes (12), signalling proteins (9) and defence proteins (4), and a considerable number of identified spots (16) were proteins of unknown function (Fig. 6C). Principal Component Analysis (PCA) performed on these differentially changed spots (Fig. 7A) indicated the presence of six clusters based on similarity of expression profiles (Fig. 7B). Proteins in Cluster A (decreased in GA_3 and GA_3 + TN aleurone samples and increased in GA_3 + HS samples) and in Cluster B (increased in GA_3 + HS samples) were mainly proteins involved in primary metabolism such as glycolysis, citric acid cycle and amino acid biosynthesis. Remarkably, 11 out of 22 spots in Cluster C with lower abundance in GA_3 + HS samples, contain small HSPs. In Cluster D, the spots from the three samples incubated with GA_3 decreased in volume when compared with untreated samples; it is also very interesting that 38% of the identifications in Cluster D were proteins of unknown function. Clusters C and E grouped protein spots whose profiles dramatically decreased in HS samples when compared with GA_3 and GA_3 + TN treated samples. Proteins in 28 spots in Cluster E belong to primary metabolism, indicating heat-induced protein degradation; remarkably 3 AMY1 and 3 AMY2 spots were assigned to Cluster E, suggesting selective destabilization of secretory protein mRNA by heat stress. Finally, spots showing increases both in GA_3 + TN and GA_3 + HS samples were grouped in the small Cluster F, including several late embryogenesis abundant proteins (LEA) and protein disulphide isomerase (PDI). Furthermore, among the six clusters there were 19 protein spots suspected to be protein fragments due to a clear discordance between theoretical and experimental pI/M_r (Supplemental Table S3). In accordance with this their sequence coverages obtained from Mascot search were confined to the N- or C-terminal part of the protein. Curiously, 17 of these 19 protein fragments belong to Cluster E.

Identification of *N*-Glycosylated Proteins from Complex Protein Samples

An *N*-glycoproteome analysis was performed in both intracellular and extracellular fractions from aleurone layers in order to provide the first overview of glycosylation sites in barley proteomes. Enrichment of glycopeptides from trypsin-digested protein

extracts was carried out by hydrophilic interaction liquid chromatography (HILIC), followed by deglycosylation using PNGase A in the presence of $H_2O^{18}O$. During the deglycosylation reaction the formerly glycosylated asparagine residues undergo deamidation and ^{18}O is incorporated into the carboxylate groups of the resulting aspartate sidechains. Deglycosylated asparagines can thus be destinguished from residues that were deamidated prior to the PNGase A reaction. Analysis by LC-MS/MS on a Q-Exactive Orbitrap enabled determination of 73 independent glycopeptides in 65 glycoproteins (Table I; Supplemental Table S4). Of these, 57 glycoproteins were found in the extracellular fraction, whereas 36 glycoproteins were identified in the intracellular samples. A total of 28 proteins were found in both fractions. All the glycoproteins were predicted to follow the secretory pathway according to the signal peptide cleavage site prediction (Supplemental Table S4).

DISCUSSION

By application of two stresses that affect protein secretion through independent mechanisms, new information about the plant protein secretion system in GA_3-induced aleurone layers has been obtained.

A clear effect of GA_3 on H_2O_2 production and cell death was confirmed (Bethke and Jones, 2001; Ishibashi et al., 2012). Moreover lipid peroxidation, as indicative of cell damage in response to several stresses (Bailly et al., 2004; Barba-Espin et al., 2011), was enhanced by GA_3 (Fig. 2A). A high TN concentration (20 μg mL^{-1}) provoked further H_2O_2 accumulation and cell death, which is probably due to the oxidative damage enhanced in these samples. The inhibition of protein folding and blocking of secretion provoked by high TN concentrations seemed to be incompatible with cellular function, whereas the response triggered by lower TN concentrations may still allow recovery of the ER function.

Overall, GA_3 induced protein secretion from aleurone layers as shown (Finnie et al., 2011). Half of the observed spots decreased in abundance under one or both of the stress treatments. Most of these contained proteases whereas the majority of the α-amylase spots remained at similar levels to unstressed (GA_3-induced) aleurone layers (Fig. 4 and 5; Supplemental Table S2). Cysteine proteinases are the most abundant proteases secreted by barley aleurone layers in response to GA_3 (Zhang and Jones, 1995; Martínez and Díaz, 2008), and the synthesis of several family members remains

high during the first 36 h of GA_3-incubation (Koehler and Tuan-Hua, 1990). By contrast, α-amylase secretion peaks much earlier during the first 24 h of incubation (Bethke et al., 1997). Thus, the difference in the responses of proteases and α-amylase to the treatments could be due to more significant secretion of cysteine proteases than α-amylases during the 4 h period of heat shock.

In the intracellular proteomes, 178 spots changed abundances, and were assigned to six clusters (Fig. 6 and 7). The overall effect of TN and HS treatments on the intracellular spot patterns reflects a differential protein response; HS provoked major changes in the protein patterns when compared with treatments with GA_3 alone and TN, as expected due to suppression of secretory protein synthesis (Brodl and Ho, 1991; Spiess et al., 1999) and the more selective mechanism of TN effect on the ER lumen. Considering the short incubation time at 40ºC, the major changes occurring in HS samples suggest changes in protein turnover. In fact, the most dramatic variations between GA_3 and stress treatments correspond to spots either decreasing or increasing in HS samples (Fig. 7; Supplementary Table S3). Complementary expression profiles are observed for Clusters A and E (Fig. 7), reflected by several full-length proteins in Cluster A (aconitate hydratase, HSP70, enolase, aldose reductase, 6-phosphogluconate dehydrogenase, phosphoenolpyruvate carboxylase and malic enzyme) being present in Cluster E as probable fragments.

The aleurone layer contains large reserves of oil, and GA_3 stimulates both fatty acid β-oxidation and the glyoxylate cycle, for the production of sucrose from storage lipids and to produce carbon skeletons for production of secreted hydrolases (Doig et al., 1975; Eastmond and Jones, 2005). Glyoxylate cycle enzymes isocitrate lyase and aconitate hydratase were identified in 12 and eight spots, respectively, a number of these containing fragments (Supplemental Table S3). Those migrating according to their predicted M_r and pI all decreased in intensity in GA_3 + TN samples, as did a predicted acyl-coenzyme-A oxidase identified in one spot (Supplemental Table S3). This suggests a down-regulation of β-oxidation and the glyoxylate cycle coupled to the decrease in production of secretory proteins.

In agreement with previous work (Finnie et al., 2011), several secretory proteins (α-amylases AMY1 and AMY2, α-*N*-arabinofuranosidase A, β-D-xylosidase, serine carboxypeptidase) could be identified in intracellular aleurone layer extracts (Supplemental Table S3). These intracellular forms are presumably *en route* in the

protein secretion pathway. Spots containing AMY1 and AMY2, α-*N*-arabinofuranosidase A, and β-D-xylosidase were decreased in the intracellular extracts subjected to TN or HS treatments. This decrease was not seen in the extracellular protein fractions (Fig. 4; Supplementary Table S2), probably reflecting that the secretory proteins are transiently present within the aleurone layers, while the extracellular fraction contains proteins accumulating over the entire incubation period. For this reason, it was not possible to observe a decrease in α-amylase in after 4 h HS by western blotting of extracellular proteins accumulated over 24 h (data not shown), although the decrease was apparent both in the intracellular fraction (Fig. 1B) and in the extracellular fraction collected solely during the 4 h period of HS (Fig 1A).

Many of the extracelular and intracellular proteins (Supplemental Table S2 and 3) migrated on gels with higher M_r than predicted, suggesting the presence of PTMs. These proteins were confirmed to have a leader peptide processing site and at least one putative *N*-glycosylation site (data not shown). Information about the PTMs on barley α-amylases is lacking, but extensive modifications are suggested by the complex pattern of spots detected in both supernatant and intracellular fractions (Bak-Jensen et al., 2007; Finnie et al., 2011).

According to Swiss-Prot database predictions, more than 50% of eukaryotic proteins are glycosylated (Apweiler et al., 1999), the majority of them being related to the secretory system. Therefore a large number of proteins are potential targets for TN action. Many studies have characterised a high number of glycoproteins from bacterial (Nothaft and Szymanski, 2010) and animal species (Bunkenborg et al., 2004; Wollscheid et al., 2009; Zielinska et al., 2010). The present work represents a major contribution towards mapping the barley glycoproteome, assigning 73 *N*-glycosylation sites in 65 proteins (Table I; Supplemental Table S4), several of which were found by 2D-based analysis either to increase (PDI, calreticulin) or decrease in response to TN or HS treatments (β-xylosidase, α-*N*-arabinofuranosidase A and serine carboxypeptidase; Fig. 7, Supplemental Table S3). According to Uni-Prot database annotations, from the 65 assigned glycoproteins, only β-xylosidase, serine carboxypeptidase, xylanase inhibitor and oxalate oxidase were previously confirmed to be *N*-glycosylated in barley.

TN treatment of GA_3-induced aleurone layers resulted in a decrease in Pro-Q® Emerald-stained glycoproteins in the extracellular fraction, including cathepsin B, cysteine proteinase C1A, two α-amylase (AMY2), two purple acid phosphatase, two

subtilisin-like serine protease and six beta-glucosidase spots (Fig. 3; Supplemental Table S1). Of these, a corresponding glycopeptide was identified in a cysteine proteinase, purple acid phosphatase, subtilisin-like serine protease and a β-glucosidase (Supplemental Table S4).

To date, no α-amylase form has been confirmed to be glycosylated. The barley seed complex 2D spot pattern obtained from a limited number of gene products (Bak-Jensen et al., 2007; Finnie et al., 2011) probably reflect PTMs. Here we showed by specific fluorescent staining that two AMY2 spots contain glycoproteins. In agreement with this, both forms migrated with an apparent molecular mass 2 kD larger than the theoretical value. Moreover, in GA_3-induced aleurone layers treated with TN, the corresponding AMY2 spots with elevated M_r decreased in intensity for both extracellular and intracellular fractions, representing the most significant intensity change in response to TN. By contrast AMY2 forms migrating according to their predicted M_r were not affected by TN. A single *N*-glycosylation site is predicted in this AMY2 gene product, however a corresponding glycopeptide was not identified. This may be due to the high hydrophilicity of this particular glycopeptide, hinding its retention on the reversed phase column prior to MS analysis. The glycosylated AMY2 forms identified here constitute a minor proportion of the total α-amylase produced by the isolated aleurone layers (Fig. 3) or in germinating grains (Bak-Jensen et al., 2007). Interestingly, the presence of glycosylated AMY2 varies among barley cultivars, since the AMY2 form with elevated M_r was only seen in some cultivars (Bak-Jensen et al., 2007). It remains to be determined whether the glycosylation affects the thermostability or kinetic properties of barley AMY2, as suggested for rice Amy1A (Terashima et al., 1994).

Plant cells seem to have a common strategy for overcoming ER stress through (i) attenuation of genes encoding secretory proteins, (ii) enhancement of protein folding activity by chaperones, such as the ER localised calreticulin or PDI, and (iii) degradation of unfolded proteins by means of the ER-associated protein degradation (ERAD) system (Travers et al., 2000; Martinez and Chrispeels, 2003; Kamauchi et al., 2005). The current study reveals components and evidence of this triple response. Five major HSP families are conserved among species (Wang et al., 2004): HSP60, HSP70; HSP90, HSP100, and small HSP (12 to 40 kD). Marked differences in proteins with chaperone functions were observed between TN and HS samples, suggesting different pathways in the induction of HSPs and ER stress response (Johnston et al., 2007).

Whereas TN treatment increased nine HSP spots (17, 26 and 70 kD), HS induced an increase in six HSP spots (17, 60 and 70 kDa) but a decrease in 17 HSP spots (mainly small HSP). Thermoprotection in the aleurone has been reported to correlate more with enhanced levels of fatty acid saturation in membrane phospholipids than with HSP expression, since HSP 70 synthesis was only slightly induced by 3 h incubation at 40ºC (Shaw and Brodl, 2003). Our results suggest an overall decrease of small HSP and an increase in certain intermediate and high M_r chaperones in HS samples; these observations indicate the differential HSP expression under HS, and highlight the secondary role of HSPs in acquiring thermotolerance in barley aleurone cells. Moreover, five spots corresponding to HSP70 fragments decreased intensity in HS samples. In fact, some other metabolic protein spots judged to be the result of proteolytic processing were decreased after HS treatment, suggesting changes in protein turnover (Fig. 7; Supplemental Table S3).

Two important ER stress-induced chaperones were found to increase in response to both TN and HS treatments: calreticulin and PDI. Family members of these ER-resident multifunctional proteins are known to be specifically activated by *IRE1/bZIP60* transcription factor, which is mediator of the ER stress response (Ye et al., 2011; Moreno et al., 2012). Calreticulin was identified in two spots with the same isoelectric point on 2D gels, but differing in molecular mass by about 2 kD, suggesting that the protein in the slower migrating spot may be *N*-glycosylated, and in the faster migrating spot a non-glycosylated form. In agreement with this hypothesis, TN caused a specific increase in intensity of the lower spot, which was the most pronounced accumulation observed in response to TN, whereas the intensity of the upper spot decreased in abundance By contrast, HS slightly reduced the level of the lower calreticulin spot, and increased the upper spot (Supplemental Table S3). Thus, HS and TN have contrasting effects on the two calreticulin-containing spots. Moreover, PDI was identified in one spot increasing upon both tunicamycin and HS treatments (Supplemental Table S3). Glycosylation of both calreticulin and PDI was confirmed by identification of the predicted glycosylation site (Supplemental Table S4). Interestingly, the glycopeptides were identified in the extracellular fraction. The significance of this is unclear, however calreticulin is known to have multiple cellular locations and in mammalian cells has been identified at the cell surface (Gardai et al., 2005; Gold et al., 2010). Further approach to gain insight on the biological role of calreticulin glycosylation is needed,

but an effect on its interaction with other proteins and subcellular localisation is expected. Two proteasome subunit spots were also induced, one of them upon TN treatment and the other after HS, both likely associated with the ERAD system.

Two 1-Cys peroxiredoxins spots differing in pI were identified (Supplemental Table S3). Besides its antioxidant role during barley germination (Stacy et al., 1999), a chaperone activity was demonstrated for 1-Cys peroxiredoxin from Chinese cabbage seeds (Kim et al., 2011). 1-Cys peroxiredoxins are susceptible to overoxidation of the active-site cysteine, which can be observed in 2D gel electrophoresis since this causes a pI-shift of the protein (Pulido et al., 2009). Chaperone activity was increased in the overoxidised form, whereas the reduced form had greater peroxidase activity (Kim et al., 2011). In our experiments, these reduced and oxidised forms likely correspond to the high and low pI spots, respectively. HS and to a lesser extent TN treatment caused an increase in the reduced form and decrease in the oxidised form, which might indicate a decrease in its chaperone activity and an enhanced antioxidant activity.

Cluster F (Supplemental Table S3) contained proteins in addition to PDI that increased in response to both stress treatments and are therefore strong candidates for having roles in ER stress mechanisms. Some of these are known to respond to various abiotic stresses, i.e. glucose and ribitol dehydrogenase (Witzel et al., 2010) and glyoxalase (Wu et al., 2013). Interestingly, a ribose-phosphate pyrophosphokinase (phosphoribosyl pyrophosphate (PRPP) synthase) was identified, which synthesises PRPP, an important intermediate in purine and pyrimidine biosynthesis (Zrenner et al., 2006) The increase in amount of this protein could suggest a need for DNA repair. A putative isopentenyl diphosphate isomerase was also identified in this cluster (Supplemental Table S3). This enzyme belongs to the mevalonate pathway, and is required for the synthesis of the polyisoprenoid lipid carrier for the glycans used in *N*-glycosylation (Jones et al., 2009). Therefore, a compensatory mechanism to restore the folding capacity of ER might be suggested, although further approaches will be required to analyse the mevalonate pathway activity under stress conditions.

Noticeably, numerous proteins induced by TN treatment including sHSPs and calreticulin (Supplemental Table S3) were also slightly upregulated in GA_3-induced compared with untreated aleurone layers. This was reflected in the PCA analysis, where grouping of the GA_3-induced and GA_3 + TN treated aleurone layers suggested similar changes in comparison to the untreated samples, in contrast to the GA_3 + HS treated

aleurone layers (Figure 7A). This suggests that the GA_3-induced aleurone layers may display early symptoms of ER stress due to the heavy load on the protein secretory machinery (Fig. 8). The experimental conditions used here were carefully chosen to avoid induction of cell death due to ER-stress, but to inflict a subtle pressure on the protein secretory system. Further investigations along these lines of evidence will explore possible links between GA_3-induced cell death and ER stress in barley aleurone layers.

Conclusions

Overall, the data presented here provide new insights into the proteomes of *in vitro* maintained aleurone layers: (i) a large number of proteins have been identified in the intracellular water-soluble fraction, many of which had previously not been identified in barley aleurone layers; (ii) we provide the first comparative study of the distinct and overlapping effects of TN and HS on a plant proteome and show that this strategy identifies key components of the protein secretory and quality control machinery, and identifies proteins with likely functions in ER-stress responses. In summary, HS entails a major degradation of the water-soluble protein fraction when compared with TN treated samples, whereas the TN effect was more selective. In particular, distinct subsets of HSPs respond divergently to HS. Common components of the ER stress response induced by the TN and HS treatments include HSPs, calreticulins, enzymes of nucleotide metabolism and isopentenyl diphosphate isomerase, and a reduction is observed in polysaccharide hydrolases and proteases in both extra- and intracellular fractions; (iii) the assignment of *N*-glycosylation sites represents a major advance in mapping the barley glycoproteome, determining the largest number of *N*-glycoproteins in this species to date. This approach provides targets for further proteomic experiments and characterization of specific protein patterns.

In conclusion, our work emphasizes new features of the barley aleurone layer, expanding its use as a suitable model for the analysis of the plant protein secretion system and the nature of plant PTMs.

MATERIAL AND METHODS

Plant Material, Experimental Design and Sample Preparation

Barley grains (*Hordeum vulgare* L. cv. Himalaya, 2003 harvest), representing a pool of individuals, were purchased from Washington State University (Pullman, WA, U.S.A.). These were used as the starting point from which replicate samples were prepared. Aleurone layers were isolated as described previously (Hynek et al., 2006). For each treatment, 100 mg of aleurone layers were incubated in 2 mL of a baseline culture medium (20 mM sodium succinate, 20 mM $CaCl_2$, pH 4.2), and incubated at room temperature with continuous gentle shaking for 24 h. When 5 μM GA_3 was added to the baseline culture, samples are referred as to GA_3-induced aleurone layers. Two additional treatments were applied to aleurone layers incubated in baseline culture + 5 μM GA_3: either TN (added from a stock in methanol) or HS (40ºC during the last 4 h of incubation).

Aleurone layers and incubation buffers containing the extracellular proteins from the above samples (untreated, GA_3-induced, GA_3 + HS and GA_3 + TN) were harvested at 24 h and samples prepared:

Intracellular protein fraction: Aleurone layers were washed four times with 2 mL of the baseline medium, frozen in liquid nitrogen, and stored at -80°C until use. Proteins were extracted from 100 mg aleurone layers in 1 mL buffer (5mM Tris-HCl, 1mM $CaCl_2$, pH 7.5; Østergaard et al., 2002) with the protease inhibitor cocktail Complete (Roche, Germany) for 30 min in a cold room. Insoluble material was pelleted by centrifugation (10 min at 13,000 rpm, 4ºC), and supernatants containing solubilized proteins were stored at -80ºC until use.

Extracellular protein fraction: The spent incubation medium from the aleurone layers was centrifuged to remove insoluble material (5 min at 13,000 rpm, 4ºC), and then supernatants containing the extracellular proteins were frozen at -80ºC until use.

At least three biological replicates were used for each analysis. In each one the whole procedure, starting from the selection of the dry seeds, was repeated.

Protein Quantification

Protein concentrations in both intracellular and extracellular protein fractions were determined using the Popov assay with bovine serum albumin as standard (Popov et al., 1975).

Determination of Cell Viability and Death

The percentage of live and dead cells was determined by double staining with FDA (2 μg mL^{-1} in 20 mM $CaCl_2$) for 15 min, followed by MM 4-64 (20 μM in 20 mM $CaCl_2$) for 2 min as described (Fath et al., 2001; Wu et al., 2010) with minor modifications. Aleurone layers were observed with a fluorescent microscope (Nikon Eclipse E1000, Nikon Instruments Inc., NY, U.S.A) using a x20 objective, and images were captured using a digital camera. Randomly selected fields from three different aleurone layers per treatment (biological replicates) were counted to determine the percentage of live cells.

Determination of Endogenous H_2O_2 and Lipid Peroxidation (TBARS)

The measurement of H_2O_2 in the aleurone layers was based on the peroxide-mediated oxidation of Fe^{2+}, followed by the reaction of Fe^{3+} with xylenol orange (Bellincampi et al., 2000). Absorbance values were calibrated to a standard curve generated with known concentrations of H_2O_2.

The extent of lipid peroxidation was assayed by determining the concentration of thiobarbituric acid-reactive substances (TBARS) as described (Hernández and Almansa, 2002).

Four biological replicates of 20 aleurone layers per treatment were used for each analysis. Volumes of samples and reactants were adapted for measuring the absorbance in 96 well microplates. Statistical significance of results was analysed according to Tukey's Test ($P < 0.05$), by using the XLSTAT software for Mac (Addinsoft, NY, U.S.A.).

Western Blotting for α-Amylase Detection

Either five μg of protein extracted from the aleurone layer (intracellular fraction) or proteins contained in 20 μl of extracellular fraction were separated on 4% to 12% Bis-Tris NuPAGE gels (Invitrogen, CA, U.S.A.) followed by electroblotting to a nitrocellulose membrane (Hybond™ - ECL, Amersham, GE-Healthcare, U.K.) before being probed with rabbit polyclonal antibodies raised against barley α-amylase (Søgaard & Svensson, 1990). Secondary goat anti-rabbit IgG antibodies conjugated to alkaline phosphatase and "BCIP/NBT" Western blotting reagent (Sigma Fast™, Sigma-Aldrich,

MO, U.S.A) were used to detect binding. Three biological replicates were performed. For each one, a gel was run in parallel and stained with silver nitrate (Heukeshoven and Dernick, 1988). For analysis of the effect of heat shock on the extracellular α-amylase, the incubation buffer was replaced at the onset of the heat shock, allowing harvest of the proteins released only during the final 4 h of incubation. To enable sufficient protein to be loaded on the gel, it was necessary to concentrate these protein samples by precipitation (4 volumes of acetone at −20°C O/N).

2-D Gel Electrophoresis and Staining

Gel for mass spectrometry analysis. Both the secreted proteins present in 2 mL of incubation buffer and proteins from aleurone layer extractions were desalted on NAP-5 columns (GE-Healthcare) following manufacturer's instruction. Aliquots containing 140 µg of proteins were precipitated with four volumes of acetone, and applied to 18 cm pH 3 - 10 L IPG strips (Immobiline^{TM} DryStrip, GE Healthcare) for first-dimension isoelectrofocusing (Shahpiri et al., 2008). The IPG strips were equilibrated (Finnie et al., 2002) and the second dimension was run on Excel 12% -14% gradient gel (GE Healthcare) using a Multiphor II (GE Healthcare; Shahpiri et al., 2008). Proteins were visualized by colloidal Coomassie brilliant blue staining (Rabilloud, 2000). At least three biological replicates were performed for each of the four treatments (four replicates were performed for GA_3 + HS (intracellular and extracellular fractions) and for the GA_3 extracellular fraction). After scanning, gel images were analysed using the Progenesis SameSpots software (version 4.5.4325.33621, Nonlinear Dynamics, UK) according to O'Gorman et al. (2007). Statistical comparison of the spot intensity among the four treatments, based on one-way ANOVA ($p < 0.05$) and a fold change greater than 1.5 (calculated as the difference between the lowest and the highest mean normalised spot volume) was verified using the false discovery rate (FDR) analysis ($q < 0.05$).

Gel Glycoprotein Staining and Fluorescence Visualization. Secreted proteins from GA_3 and GA_3 + 5 µg mL^{-1} TN samples (100 µg) were precipitated and run into 11 cm pH 3 - 10 L IPG strips as described above. The IPG strips were then trimmed to 7 cm covering pH range from 4 to 8.5, and the second dimension SDS-PAGE (NuPAGE® Novex 4-12% Bis-Tris ZOOM® Gel, Invitrogen) was performed in the XCell SureLock™ Mini-Cell system (Invitrogen). Gels were stained with Pro-Q® Emerald 300 Glycoprotein

Gel Kit (Molecular Probes, Eugene, USA) according to the manufacturer's instructions. Three biological replicates were performed for each of the two treatments. The fluorescent spots were visualized using UVP BioSpectrum® Image System (AH diagnostics, Denmark) with transillumination at 302-nm. Images of gels were obtained using a 485–655 nm filter. After imaging, the gels were stained with colloidal Coomassie brilliant blue.

In-Gel Digestion and Protein Identification by MALDI TOF/TOF

Spots of interest were excised from 2D SDS-PAGE gels and digested as described by Shevchenko et al. (1996). After soaking with trypsin (modified porcine trypsin, Promega, Madison, WI) the gel pieces were covered with 70 µL of 10 mM $(NH_4)HCO_3$ and incubated at 37°C o/n. The supernatant containing tryptic peptides was transferred to a clean tube and 2 µL were placed onto a MALDI anchor target, The sample was dried at room temperature and covered with 1 µL of 5 µg μL^{-1} α-cyano-hydroxycinnamic acid in 70% (v/v) acetonitrile (ACN), 0.1% (v/v) trifluoroacetic acid (TFA), dried again and washed with 2 µL of 0.5% (v/v) TFA. An Ultraflex II MALDI-TOF-TOF mass spectrometer (Bruker-Daltonics, Bruker, Germany) was used in positive ion reflector mode and spectra were analysed using FlexAnalysis software (Bruker-Daltonics). Internal calibration was carried out with trypsin autolysis products (m/z 842.51 and m/z 2211.10).

The peptide masses and sequences obtained were matched to proteins with an in-house MASCOT server (Version 2.2.04), searching in the MIPS barley genome database containing high confidence gene models (ftp://ftpmips.helmholtz-muenchen.de/plants/barley/public_data/) Search parameters were: allowed modifications, oxidation of Met (variable) and carbamidomethylation of Cys (static); monoisotopic mass tolerance, 40 ppm; allowed missed cleavages, 1.

Positive protein identifications were considered when one of the following requirements was met: three independent peptides where one of them was confirmed by MS/MS fragment ion fingerprinting; or two independent peptides, both confirmed by MS/MS. For more details see Supplemental Tables S1, S2 and S3.

Identification of *N*-Glycosylated Proteins from Complex Samples and Assignment of Their Glycosylation Sites

In-Solution Digestion and Hydrophilic Interaction Liquid Chromatography (HILIC). Protein samples from both intracellular and extracellular fractions were digested in-solution as described (Hägglund et al., 2004) and applied to microcolumns containing a combination of ZIC-HILIC resin and cotton. A small piece of cotton wool (approximately 200 µg) taken from a cotton wool pad was pushed into a 20 µL GELoader tip (Eppendorf, Hamburg, Germany) and ZIC-HILIC resin (kind gift of Sequant, Umeå, Sweden; particle size 10 μm) was packed on top of the cotton wool. The microcolumns were equilibrated with 40 µL of 80% (v/v) ACN, 0.5% (v/v) formic acid (FA). In-solution trypsin digested protein samples (20 µg) were dried (SPD 1010 Speed Vac, Thermo Scientific, CA, U.S.A), redissolved in 40 µL 80% (v/v) ACN, 0.5% (v/v) FA and applied onto the equilibrated HILIC micro-columns. The column was washed twice with 80% (v/v) ACN, 0.5% (v/v) formic acid (FA) followed by elution of bound peptides in 0.5% (v/v) FA.

Endoglycosidase Digestion and Protein Identification by Q-Exactive LC-MS/MS. Deglycosylation of the eluted peptides was performed in 100 mM ammonium acetate prepared in $H_2{}^{18}O$-water for 24 h at 37°C with 0.1 mU *N*-glycosidase A (PNGase A, Roche). To minimize ^{18}O-isotope labeling at peptide C-termini, the deglycosylated samples were dried, redissolved in 100 mM ammonium acetate prepared in $H_2{}^{16}O$-water and incubated for 24 h at 37°C in the presence of 0.2 µg sequencing grade modified trypsin (Promega, WI, U.S.A.). The peptides were desalted on a StageTip C18-reversed phase column (Thermo Scientific) and separated on an EASY-Spray column (Pepmap 3 µm, C18 15 cm x 75 µm) coupled on-line to a Q Exactive^{TM} orbitrap (Thermo Scientific,). The separation was performed using a flow rate of 300 nL min^{-1} and a gradient of solvents A (0.1% FA) and B (80% ACN, 0.1% FA) as follows: 5–40% B for 40 min; 40–100% B for 5 min; 100% B for 10 min. The LC-MS data were processed using Proteome Discoverer Software (version 1.3, Thermo Scientific), then searched using MASCOT (Version 2.2.04) against the MIPS barley genome database (as above). MS and MS/MS accuracy were 10 ppm and 20 mmu, respectively. In addition to oxidation of methionine, variable modifications were set to include deamidation of asparagine or glutamine with incorporation of ^{16}O (asparagine/glutamine + 0.98 D), ^{18}O (asparagine/glutamine + 2.99 D), and the incorporation of one or two ^{18}O at the peptide

C-terminus. Carbamidomethyl cysteine was set as a fixed modification. In order to be accepted as a glycopeptide, the deamidated residue containing ^{18}O atoms had to be situated within a consensus sequence for *N*-glycoslation (NXS/T).

SUPPLEMENTAL DATA

The following materials are available in the online version of this article:

Supplemental Table S1. Identification of candidate glycoproteins by Pro-Q® Emerald staining in extracellular fractions of GA_3-induced aleurone layers

Supplemental Table S2. Proteins identified in extracellular fractions of 24 h-incubated aleurone layers

Supplemental Table S3. Proteins identified in intracellular water-soluble fractions of 24 h-incubated aleurone layers

Supplemental Table S4. Identification of *N*-glycosylated proteins from complex samples and assignment of their glycosylation sites

ACKNOWLEDGEMENTS

The authors thank Birgit Andersen for excellent laboratory assistance and mass spectrometry analysis, and Martin Kogle for help with microscopy experiments. Susanne Jacobsen is also thanked for valuable discussion and support.

LITERATURE CITED

Agrawal GK, Jwa NS, Lebrun MH, Job D, Rakwal R (2010) Plant secretome: unlocking secrets of the secreted proteins. Proteomics **10**: 799-827

Alexandersson E, Ali A, Resjö S, Andreasson E (2013) Plant secretome proteomics. Front Plant Sci. **4**: 1-6

Apweiler R, Hermjakob H, Sharon N (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta **1473**: 4-8

Bailly C (2004) Active oxygen species and antioxidants in seed biology. Seed Sci Res **14**: 93-107

Bailly C, El-Maarouf-Bouteau H, Corbineau F (2008) From intracellular signaling networks to cell death: the dual role of reactive oxygen species in seed physiology. C R Biol **331**: 806-814

Bak-Jensen KS, Laugesen S, Ostergaard O, Finnie C, Roepstorff P, Svensson B (2007) Spatio-temporal profiling and degradation of α-amylase isozymes during barley seed germination. FEBS J **274**: 2552-2565

Barba-Espín G, Diaz-Vivancos P, Job D, Belghazi M, Job C, Hernández JA (2011) Understanding the role of H_2O_2 during pea seed germination: a combined proteomic and hormone profiling approach. Plant Cell Environ **34**: 1907-1919

Belanger FC, Brodl MR, and Ho TH (1986) Heat shock causes destabilization of specific mRNAs and destruction of endoplasmic reticulum in barley aleurone cells. Proc Natl Acad Sci U S A **83**: 1354-1358

Bellincampi D, Dipierro N, Salvi G, Cervone F, De Lorenzo G (2000) Extracellular H_2O_2 induced by oligogalacturonides is not involved in the inhibition of the auxine-regulated rolB gene expression in tobacco leaf explants. Plant Physiol **122**: 1379-1385

Bethke P, Schuurink R, Jones RL (1997) Hormonal signalling in cereal aleurone. J Exp Bot **48**: 1337-1356

Bethke PC, Jones RL (2001) Cell death of barley aleurone protoplasts is mediated by reactive oxygen species. Plant J **25**: 19-29

Bond U, Schlesinger MJ (1986) The chicken ubiquitin gene contains a heat shock promoter and expresses an unstable mRNA in heat-shocked cells. Mol Cell Biol **6**: 4602-4610

Bønsager BC, Finnie C, Roepstorff P, Svensson B (2007) Spatio-temporal changes in germination and radical elongation of barley seeds tracked by proteome analysis of dissected embryo, aleurone layer, and endosperm tissues. Proteomics **7**: 4528-4540

Brodl MR, Ho THD (1991) Heat shock causes selective desta- bilization of secretory protein mRNAs in barley aleurone cells. Plant Physiol **96**: 1048-1052

Bunkenborg J, Pilch B J, Podtelejnikov AV, Wisniewski JR (2004) Screening for *N*-glycosylated proteins by liquid chromatography mass spectrometry. Proteomics **4**: 454-465

Bush DS, Cornejo MJ, Huang C-N, Jones RL (1986) Ca^{2+} stimulated secretion of a-amylase during development in barley aleurone protoplasts. Plant Phys **82**: 566-74

Chrispeels MJ, Varner JE (1967) Gibberellic acid-enhanced synthesis and release of α-amylase and ribonuclease by isolated barley and aleurone layers. Plant Physiol **42**: 398-406

Craig EA, Weissman JS, Horwich AL (1994) Heat shock proteins and molecular chaperones: mediators of protein conformation and turnover in the cell. Cell **78**: 365-372

De Wilde K, De Buck S, Vanneste K, Depicker A (2013) Recombinant antibody production in Arabidopsis seeds triggers an unfolded protein response. Plant Physiol **161**:1021-1033

Denecke J, Carlsson LE, Vidal S, Höglund AS, Ek B, van Zeijl MJ, Sinjorgo KM, Palva ET (1995) The tobacco homolog of mammalian calreticulin is present in protein complexes *in vivo*. Plant Cell **7**: 391-406

Doig RI, Colborne AJ, Morris G, Laidman DL (1975) The induction of glyoxysomal enzyme activities in the aleurone cells of germinating wheat. J Exp Bot **26**: 387-398

Eastmond PJ, Jones RL (2005) Hormonal regulation of gluconeogenesis in cereal aleurone is strongly cultivar dependent and gibberellin action involves SLENDER1 but not GAMYB. Plant J **44**: 483-493

Erlendsson LS, Muench MO, Hellman U, Hrafnkelsdóttir SM, Jonsson A, Balmer Y, Mäntylä E, Örvar BL (2010) Barley as a green factory for the production of functional Flt3 ligand. Biotechnol J **5**:163-171

Fath A, Bethke PC, Jones RL (2001) Enzymes that scavenge reactive oxygen species are down-regulated prior to gibberellic acid-induced programmed cell death in barley aleurone. Plant Physiol **126**: 156-166

Finnie C, Andersen B, Shahpiri A, Svensson B (2011) Proteomes of the barley aleurone layer: A model system for plant signalling and protein secretion. Proteomics **11**: 1595-1605

Finnie C, Svensson B (2003) Feasibility study of a tissue-specific approach to barley proteome analysis: aleurone layer, endosperm, embryo and single seeds. J Cereal Sci **38**: 217-227

Fitchette AC, Dinh OT, Faye L, Bardor M (2007) Plant proteomics and glycosylation. Methods Mol Biol **355**: 317-342

Gardai SJ, McPhillips KA, Frasch SC, Janssen WJ, Starefeldt A, Murphy-Ullrich JE, Bratton DL, Oldenborg PA, Michalak M, Henson PM (2005) Cell-surface calreticulin initiates clearance of viable or apoptotic cells through trans-activation of LRP on the phagocyte. Cell **123**: 321-34

Gold LI, Eggleton P, Sweetwyne MT, Van Duyn LB, Greives MR, Naylor SM, Michalak M, Murphy-Ullrich JE (2010) Calreticulin: non-endoplasmic reticulum functions in physiology and disease. FASEB J **24**: 665-683

Hägglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P (2004) A new strategy for identification of *N*-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. J Proteome Res **3**: 556-566

Hägglund P, Bunkenborg J, Yang F, Harder LM, Finnie C, Svensson B (2010) Identification of thioredoxin target disulfides in proteins released from barley aleurone layers. J Proteomics **73**: 1133-1136

Hernández JA, Almansa MS (2002) Short-term effects of salt stress on antioxidant systems and leaf water relations of pea leaves. Physiol Plant **115**: 251-257

Heukeshoven J, Dernick R (1988) Improved silver staining procedure for fast staining in PhastSystem Development Unit. I. Staining of sodium dodecyl sulfate gels. Electrophoresis **9**: 28-32

Hynek R, Svensson B, Jensen ON, Barkholt V, Finnie C (2006) Enrichment and identification of integral membrane proteins from barley aleurone layers by reversed-phase chromatography, SDS-PAGE and LC-MS/MS. J Proteome Res **5**: 3105-3113

Ishibashi Y, Tawaratsumida T, Kondo K, Kasa S, Sakamoto M, Aoki N, Zheng SH, Yuasa T, Iwaya-Inoue M (2012) Reactive oxygen species are involved in gibberellin/abscisic acid signaling in barley aleurone cells. Plant Physiol **158**: 1705-1714

Jamet E, Albenne C, Boudart G, Irshad M, Canut H, Pont-Lezica R (2008) Recent advances in plant cell wall proteomics. Proteomics **8**: 893-908

Jerkovic A, Kriegel AM, Bradner JR, Atwell BJ, Roberts TH, Willows RD (2010) Strategic distribution of protective proteins within bran layers of wheat protects the nutrient-rich endosperm. Plant Physiol **152**: 1459-1470

Johnston MK, Jacob NP, Brodl MR (2007) Heat shock-induced changes in lipid and protein metabolism in the endoplasmic reticulum of barley aleurone layers. Plant Cell Physiol **48**: 31-41

Jones MB, Rosenberg JN, Betenbaugh MJ, Krag SS (2009) Structure and synthesis of polyisoprenoids used in *N*-glycosylation across the three domains of life. Biochim Biophys Acta **1790**: 485-494

Jones RL, Jacobsen JV (1991) Regulation of synthesis and transport of secreted proteins in cereal aleurone. Int Rev Cyt **126**: 49-88

Kamauchi S, Nakatani H, Nakano C, Urade R (2005) Gene expression in response to endoplasmic reticulum stress in Arabidopsis thaliana. FEBS J **272**: 3461-3476

Kim SY, Paeng SK, Nawkar GM, Maibam P, Lee ES, Kim KS, Lee DH, Park DJ, Kang SB, Kim MR, Lee JH, Kim YH, Kim WY, Kang CH (2011) The 1-Cys peroxiredoxin, a regulator of seed dormancy, functions as a molecular chaperone under oxidative stress conditions. Plant Sci **181**: 119-124

Koehler SM, Tuan-Hua D (1990) Ho2 Hormonal Regulation, Processing, and Secretion of Cysteine Proteinases in Barley Aleurone Layers. Plant Cell **2**: 769-783

Laubin B, Lullien-Pellerin V, Nadaud I, Gaillard-Martiniec B, Chambond C, Branlard G (2008) Isolation of the wheat aleurone layer for 2D electrophoresis and proteomics analysis. J Cereal Sci **48**: 709-714

Martínez IM, Chrispeels MJ (2003) Genomic Analysis of the Unfolded Protein Response in Arabidopsis Shows Its Connection to Important Cellular Processes. Plant Cell **15**: 561-576

Martinez M, Diaz I (2008) The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. BMC Evol Biol **8**: 198

Maya-Ampudia V, Bernal-Lugo I (2006) Redox-sensitive target detection in gibberellic acid-induced barley aleurone layer. Free Radic Biol Med **40**: 1362-1368

Melo-Braga MN, Verano-Braga T, León IR, Antonacci D, Nogueira FC, Thelen JJ, Larsen MR, Palmisano G (2012) Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botrana* infection. Mol Cell Proteomics **11**: 945-956

Meziani S, Nadaud I, Gaillard-Martinie B, Chambon C, Benali M, Branlard G (2012) Proteomic analysis of the mature kernel aleurone layer in common and durum wheat **55**: 323-330

Minic Z, Jamet E, Négroni L, Arsene der Garabedian P, Zivy M, Jouanin L (2007) A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. J Exp Bot **58**: 2503-2512

Moreno AA, Mukhtar MS, Blanco F, Boatwright JL, Moreno I, Jordan MR, Chen Y, Brandizzi F, Dong X, Orellana A, Pajerowska-Mukhtar KM (2012). *IRE1/bZIP60*-mediated unfolded protein response plays distinct roles in plant immunity and abiotic stress responses. PLoS ONE **7**:1-24

Nagashima Y, Mishiba K, Suzuki E, Shimada Y, Iwata Y, Koizumi N (2011) Arabidopsis IRE1 catalyses unconventional splicing of bZIP60 mRNA to produce the active transcription factor. Sci Rep **1**: 29-38

Noh SJ, Kwon CS, Oh DH, Moon JS, Chung WI (2003). Expression of an evolutionarily distinct novel BiP gene during the unfolded protein response in *Arabidopsis thaliana*. Gene **311**: 81-91

Nothaft H, Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. Nat Rev Microbiol **8**: 765-778

O'Gorman M, Beauvallet C, Lepercq P, David O, Seksik P, Beaugerie L, Dore J, Martin P, Bogard P, Juste C (2007) An investigation into Crohn's disease using the Progenesis SameSpots analysis platform. 24th Journée Françaises de Spectrométrie de Masse, PAU 2007.

Østergaard O, Melchior S, Roepstorff P, Svensson B (2002) Initial proteome analysis of mature barley seeds and malt. Proteomics **2**: 733-799

Palmisano G, Antonacci D, Larsen MR (2010) Glycoproteomic profile in wine: a 'sweet' molecular renaissance. J Proteome Res **9**: 6148-6159

Popov N, Schmitt M, Schulzeck S, Matthies H (1975) Reliable micro method for determination of the protein content in tissue homogenates. Acta Biol Med Ger **34**: 1441-1446

Pulido P, Dominguez F, Cejudo FJ (2009) A hydrogen peroxide detoxification system in the nucleus of wheat seed cells: protection or signaling role? Plant Signal Behav **4**: 23-25

Rabilloud T (2000) Detecting proteins separated by 2-D gel electrophoresis. Anal Chem **72**: 48-55

Reis P, Rosado G, Silva L, Oliveira L, Oliveira L, Costa M, Alvim F, Fontes E (2011) The binding protein BiP attenuates stress-induced cell death in soybean via modulation of the N-rich protein-mediated signaling pathway. Plant Physiol **157**: 1853-1865

Shahpiri A, Svensson B, Finnie C (2008) The NADPH-dependent thioredoxin reductase/thioredoxin system in germinating barley seeds: gene expression, protein profiles, and interactions between isoforms of thioredoxin h and thioredoxin reductase. Plant Physiol **146**: 789-799

Shaw AE, Brodl MR (2003) Heat shock response of warm-incubated barley aleurone layers. Am J Bot **90**: 40-48

Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem **68**: 850-858

Song W, Mentink RA, Henquet MG, Cordewener JH, van Dijk AD, Bosch D, America AH, van der Krol AR (2013) *N*-glycan occupancy of Arabidopsis *N*-glycoproteins. J Proteomics **doi**: 10.1016/j.jprot.2013.07.032

Søgaard M, Svensson B (1990) Expression of cDNAs encoding barley alpha-amylase 1 and 2 in yeast and characterization of the secreted proteins. Gene **94**: 173-179

Spiess C, Beil A, Ehrmann, M (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell **97**: 339-347

Stacy RAP, Nordeng TW, Culianez-Macia FA, Aalen RB (1999) The dormancy-related peroxiredoxin anti-oxidant, PER1, is localized to the nucleus of barley embryo and aleurone cells. Plant J **19**: 1-8

Takeichi T, Takeuchi J, Kaneko T, Kawasaki S (1998) Purification and characterization of a galactose-rich basic glycoprotein in tobacco. Plant Physiol **116**: 477-483

Tasleem-Tahir A, Nadaud I, Girousse C, Martre P, Marion D, Branlard G (2011) Proteomic analysis of peripheral layers during wheat (*Triticum aestivum* L.) grain development. Proteomics **11**: 371-379

Terashima M, Kubo A, Suzawa M, Itoh Y, Katoh S (1994) The roles of the N-linked carbohydrate chain of rice α-amylase in thermostability and enzyme kinetics. Eur J Biochem **226**: 249-254

Thannhauser TW, Shen M, Sherwood R, Howe K, Fish T, Yang Y, Chen W, Zhang S (2013) A workflow for large-scale empirical identification of cell wall N-linked glycoproteins of tomato (*Solanum lycopersicum*) fruit by tandem mass spectrometry. Electrophoresis **34**: 2417-2431

Travers K, Patil1 CK, Wodicka L, Lockhart DJ, Weissman JS, Walter P (2000) Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. Cell **101**: 249-258

Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci **9**: 244-252

Witzel K, Weidner A, Surabhi G-K, Varshney RK, Kunze G, Buck-Sorlin GH, Börner A, Mock H-P (2010) Comparative analysis of the grain proteome fraction in barley genotypes with contrasting salinity tolerance during germination. Plant Cell Env **33**: 211-222

Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, Bibel M, Schiess R, Aebersold R, Watts JD (2009) Mass-spectrometric identification and relative quantification of *N*-linked cell surface glycoproteins. Nat Biotechnol **27**: 378-386

Wu C, Ma CQ, Pan Y, Gong SL, Zhao CX, Chen SX, Li HY (2013) Sugar beet M14 glyoxalase I gene can enhance plant tolerance to abiotic stresses. J Plant Res **126**: 415-425

Wu M, Huang J, Xu S, Ling T, Xie Y, Shen W (2011) Haem oxygenase delays programmed cell death in wheat aleurone layers by modulation of hydrogen peroxide metabolism. J Exp Bot **62**: 235-248

Ye C, Dickman MB, Whitham SA, Payton M, Verchot J (2007) The unfolded protein response is triggered by a plant viral movement protein. Plant Physiol **156**: 741-755

Zielinska DF, Gnad F, Wisniewski J R, Mann M (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. Cell **141**: 897-907

Zhang N, Jones BL (1995) Characterization of germinated barley endoproteolytic enzymes by two dimensional gel electrophoresis. J Cereal Sci **21**: 145-153

Zhang S, Sherwood RW, Yang Y, Fish T, Chen W, McCardle JA, Jones RM, Yusibov V, May ER, Rose JK, Thannhauser TW (2012) Comparative characterization of the glycosylation profiles of an influenza hemagglutinin produced in plant and insect hosts. Proteomics **12**: 1269-1288

Zrenner R, Stitt M, Sonnewald U, Boldt R (2006) Pyrimidine and purine biosynthesis and degradation in plants. Annu Rev Plant Biol **57**: 80-836

FIGURE LEGENDS

Figure 1. Western blots probing α-amylase in the extracellular (A) and intracellular water-soluble protein fractions (B) of *in vitro* incubated aleurone layers. Equal volumes of extracellular protein fraction (A) and equal amounts of total protein (B) were applied into SDS-PAGE. For each blot, a parallel protein gel was stained with silver nitrate. Bands corresponding to the α-amylase (45 kD) are indicated by arrows. Densitometric quantification of Western blot signals was performed using ImageJ Software (National Institutes of Health), and expressed as the average of three biological replicates plus the standard error (bar charts). U: untreated; TN: tunicamycin; HS: 4 h heat shock.

Figure 2. Enhanced cell death in GA_3-induced aleurone layers correlates with higher endogenous H_2O_2 and lipid peroxidation contents. (A) Hydrogen peroxide and thiobarbituric reactive substances were measured in aleurone layer extracts. (B) The

vital staining was performed in intact aleurone layers and randomly selected fields were counted to determine the percentage of dead cells. Different letters indicate statistical significance according to Tukey's test ($P < 0.05$). TBARS: thiobarbituric reactive substances; FW: fresh weight; U: untreated; TN: tunicamycin; HS: heat shock.

Figure 3. Fluorescent glycoprotein staining of extracellular proteins**.** (A) One representative glycoprotein-stained gel (1 and 3) and corresponding colloidal Coomassie blue-stained gel (2 and 4) covering the pH range 4–8.5 is shown for GA_3-induced (gels 1 and 2) and for GA_3 + 5µg ml^{-1} tunicamycin-treated (TN; gels 3 and 4) aleurone layers. CandyCane molecular size markers are indicated. (B) View of the GA_3 reference gel framed in (A) showing 53 intense fluorescent spots present in all GA_3 replicates. Twenty-three of them (numbered spots) correspond to spots absent in GA_3 + TN aleurone layers, of which 14 were identified by mass spectrometry. (For details about protein identifications, see Supplemental Table S1).

Figure 4. Extracellular protein profiles of barley aleurone layer extracts**.** (A) One representative colloidal Coomassie blue-stained gel covering the pH range 3–10 is shown for each treatment. Molecular size markers are indicated. (B) Close-up view of the reference gel showing 22 identified spots. (C) Functional categories of the plant proteins identified. U: untreated; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S2).

Figure 5. Principal Component Analysis (PCA) and clustering of extracellular protein spots. (A) PCA was performed on the 22 identified spots. Biological replicates are grouped by circles and spots are indicated by numbers. (B) Expression profiles of protein spots from PCA analysis were grouped in 3 clusters. Normalised volume for each spot is expressed relative to a reference gel (GA_3), in which all spot volumes are by default set to 1. Corresponding functional classifications are indicated (PH: polysaccharide hydrolase; PR: protease; SR: stress response; and DF: defence). U: untreated; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S2).

Figure 6. Intracellular water-soluble protein profiles of barley aleurone layer extracts. (A) One representative colloidal Coomassie blue-stained gel covering the pH range 3–

10 is shown for each treatment. Molecular size markers are indicated. (B) Close-up view of the reference gel showing 178 identified spots. (C) Functional categories of the plant proteins identified. U: untreated; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S3).

Figure 7. Principal Component Analysis (PCA) and clustering of intracellular water-soluble protein spots. (A) PCA was performed on the 178 identified spots. Biological replicates are grouped by circles and spots are indicated by numbers. (B) Expression profiles of protein spots from PCA analysis were grouped in 6 clusters. Normalised volume for each spot is expressed relative to a reference gel (GA_3), in which all spot volumes are by default set to 1. Corresponding functional classifications are indicated (PM: primary metabolism; CH: chaperone; PH: polysaccharide hydrolase; DF: defence; DT: detoxification enzymes; SN: signalling; and UF: unknown function). U: untreated; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S3).

Figure 8. Summary of distinct and overlapping responses of GA_3-induced aleurone layers to heat shock (HS) and tunicamycin (TN) treatments. GA_3-induced aleurone layers may display early symptoms of ER stress due to the pressure on the protein secretion system and reflected by enhanced expression of proteins induced by ER stress.

TABLES

Table I. A total of 65 *N*-glycoproteins were identified from extra- and intracellular complex samples. Independent glycopeptides identified by LC-MS/MS analysis are placed in brackets [$^{(n)}$]. For more details about protein identifications, see Supplemental Table S4.

protein name	accession no	protein name	accession no
acidic endochitinase	MLOC_72826.1 [1]	lipid transfer proteins	AK369929 [1]
alpha-galactosidase	AK364576 [1]		AK367409 [1]
alpha-glucosidase	MLOC_66806.2 [2]	lysosomal alpha-mannosidase	AK364106 [2]
amino oxidase	AK353856 [1]	MtN19-like protein	MLOC_73878.1 [2]
aminoacylase-1	AK376199 [1]	pectin lyase-like superfamily	MLOC_75889.3 [1]
	MLOC_64721.1 [1]	pectinesterase	AK366148 [1]
aminopeptidase M1	AK358827 [1]	peptidyl-prolyl cis-trans isomerase	MLOC_38535.3 [1]
arabinofuranosidases A	MLOC_44256.2 [1]	peroxidase	MLOC_56862.1 [1]
	MLOC_56099.3 [1]	PLC-like phosphodiesterase	MLOC_65228.1 [1]
aspartic protease family proteins	AK376209 [1]	protein disulphide isomerase	AK370108 [1]
	MLOC_13635.1 [1]	purple acid phosphatases	AK360979 [1]
auxin induced-like protein	MLOC_57260.1 [1]		AK354837 [1]
beta-D-xylosidase	MLOC_62475.1 [1]		MLOC_34761.1 [1]
beta-fructofuranosidase	AK252358.1 [1]	receptor-like protein kinases	AK364371 [1]
beta-galactosidase	AK252929.1 [1]		MLOC_53309.2 [2]
beta-glucosidase-like hydrolase	AK374484 [1]	secretory peroxidases	AK365489 [1]
calreticulin	MLOC_67890.1 [1]		MLOC_72727.1 [1]
cathepsin F-like cysteine protease	MLOC_61862.1 [1]	serine carboxypeptidases	AK365716 [1]
chitinase	AK376513 [1]		MLOC_55542.1 [2]
early nodulin-like protein 3	MLOC_4663.1 [1]		AK372814 [1]
endonuclease	MLOC_5210.1 [1]		AK362092 [1]
epidermal growth factor-like	AK370029 [1]	subtilisin-like protease	MLOC_64136.1 [1]
GDPD enzyme	AK248297.1 [2]	thioredoxin like protein	AK359722 [2]
GDSL esterase/lipases	MLOC_75385.1 [1]	transmembrane 9 family protein	MLOC_11918.1 [1]
	MLOC_12361.1 [1]	UDP-glucose 4-epimerase	MLOC_60662.1 [1]
globulin 3	MLOC_59994.1 [1]	Unnanotated protein	MLOC_5243.1 [1]
glucan 1,3-beta-glucosidase	AK376623 [1]		MLOC_10934.1 [1]
glutamyl-tRNA amidotransferase	MLOC_60550.3 [1]	vacuolar sorting receptor	AK248808.1 [1]
GPI-anchored protein 2	MLOC_57195.1 [1]	vicilin-like antimicrobial peptide	MLOC_57363.1 [1]
IAA-amino acid hydrolase ILR1	AK368695 [1]	xylanase inhibitors	AK366716 [1]
kinase, putative	MLOC_34836.3 [1]		MLOC_26558.1 [1]
lectin-domain receptor-like kinase	MLOC_58449.1 [1]	xylose isomerase	MLOC_55107.1 [1]
leucine-rich receptor kinase	MLOC_1088.10 [1]		

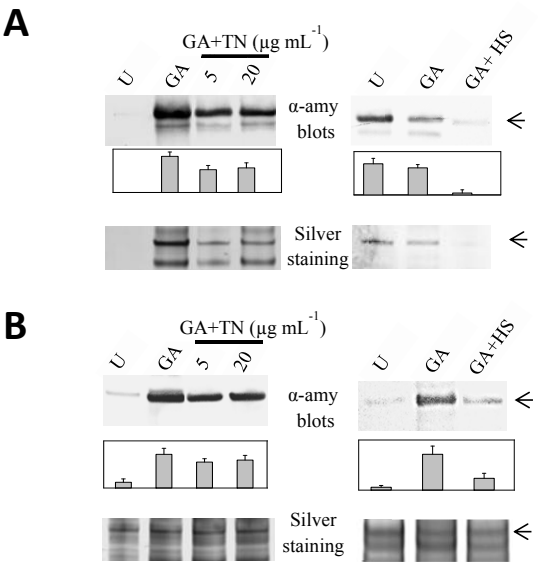


Figure 1. Western blots probing α-amylase in the extracellular (A) and intracellular water-soluble protein fractions (B) of *in vitro* incubated aleurone layers. Equal volumes of extracellular protein fraction (A) and equal amounts of total protein (B) were applied into SDS-PAGE. For each blot, a parallel protein gel was stained with silver nitrate. Bands corresponding to the α-amylase (45 kD) are indicated by arrows. Densitometric quantification of Western blot signals was performed using ImageJ Software (National Institutes of Health), and expressed as the average of three biological replicates plus the standard error (bar charts). U: untreated; GA: gibberellic acid; TN: tunicamycin; HS: 4 h heat shock.

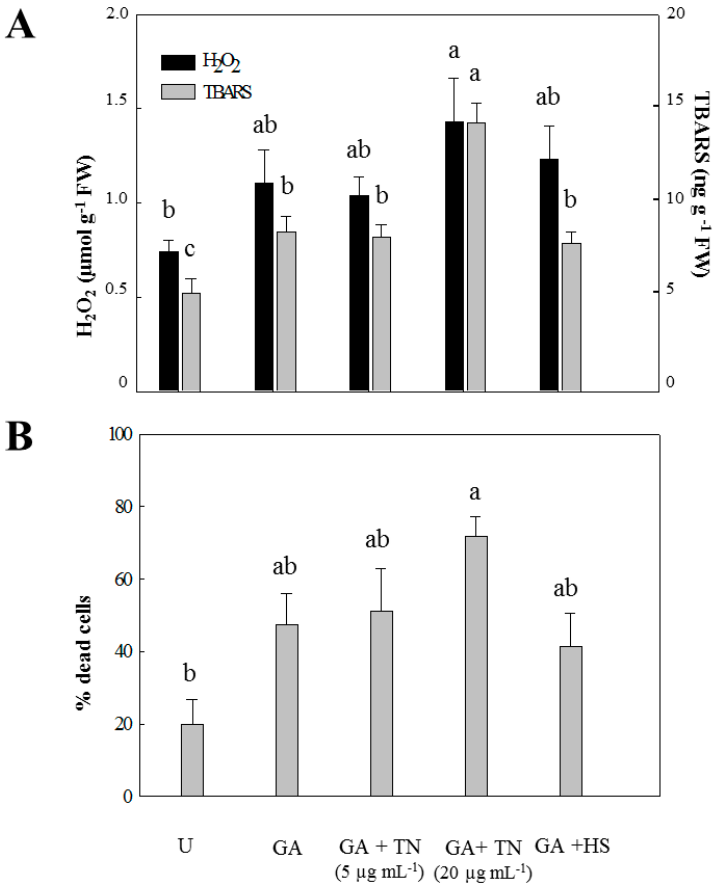


Figure 2. Enhanced cell death in gibberellic-induced aleurone layers correlates with higher endogenous H_2O_2 and lipid peroxidation contents. (A) Hydrogen peroxide and thiobarbituric reactive substances were measured in aleurone layer extracts. (B) The vital staining was performed in intact aleurone layers and randomly selected fields were counted to determine the percentage of dead cells. Different letters indicate statistical significance according to Tukey's test ($P < 0.05$). TBARS: thiobarbituric reactive substances; FW: fresh weight; U: untreated; GA: gibberellic acid; TN: tunicamycin; HS: heat shock.

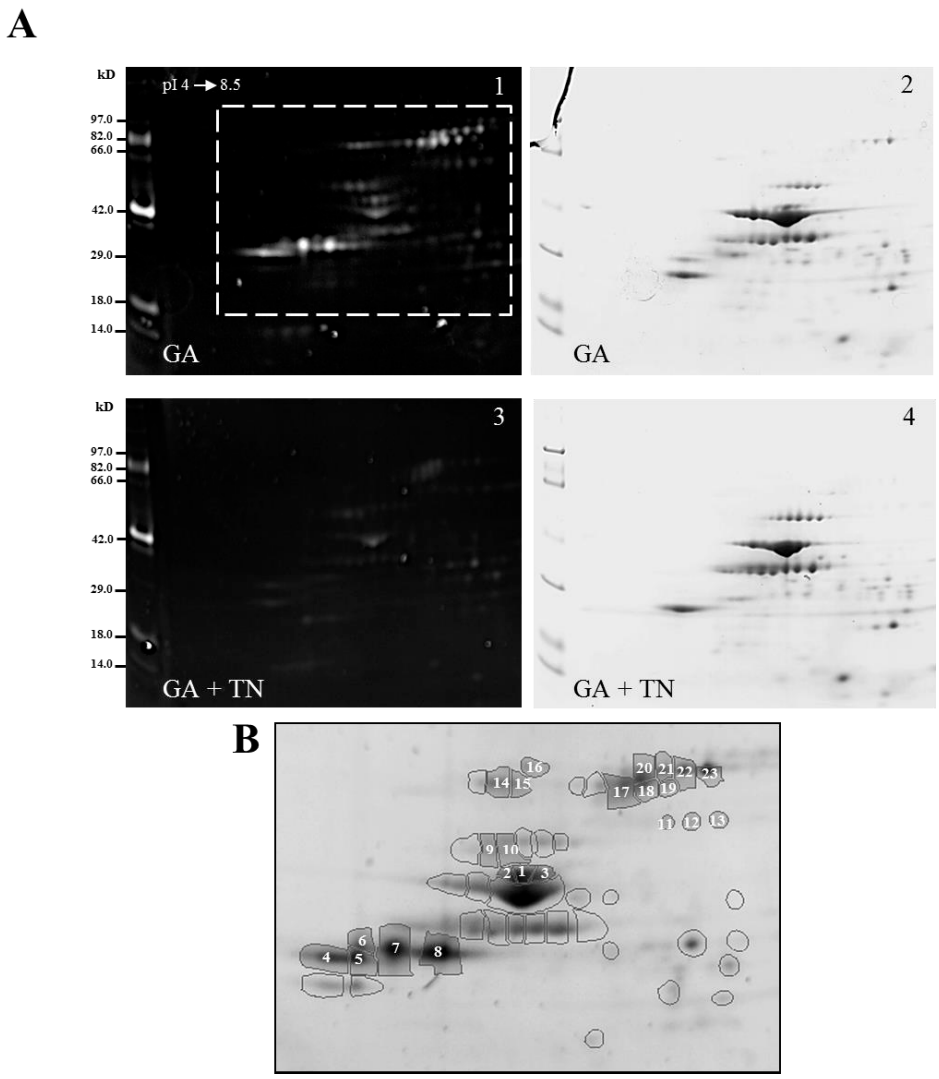


Figure 3. Fluorescent glycoprotein staining of extracellular proteins**.** (A) One representative glycoprotein-stained gel (1 and 3) and corresponding colloidal Coomassie blue-stained gel (2 and 4) covering the pH range 4–8.5 is shown for gibberellic acid-induced (GA; gels 1 and 2) and for GA + 5 µg mL^{-1} tunicamycin-treated (TN; gels 3 and 4) aleurone layers. CandyCane molecular size markers are indicated. (B) View of the GA reference gel framed in (A) showing 53 intense fluorescent spots present in all GA replicates. Twenty-three of them (numbered spots) correspond to spots absent in GA + TN aleurone layers, of which 14 were identified by mass spectrometry. (For details about protein identifications, see Supplemental Table S1).

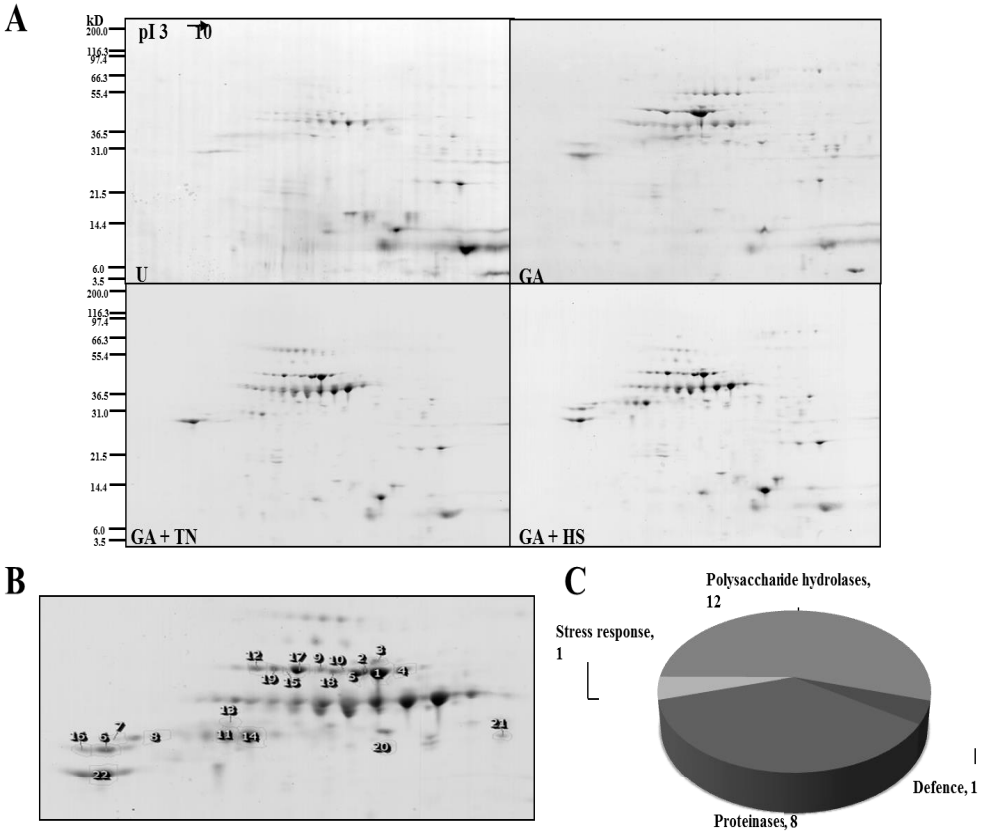


Figure 4. Extracellular protein profiles of barley aleurone layer extracts**.** (A) One representative colloidal Coomassie blue-stained gel covering the pH range 3–10 is shown for each treatment. Molecular size markers are indicated. (B) Close-up view of the reference gel showing 22 identified spots. (C) Functional categories of the plant proteins identified. U: untreated; GA: gibberellic acid; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S2).

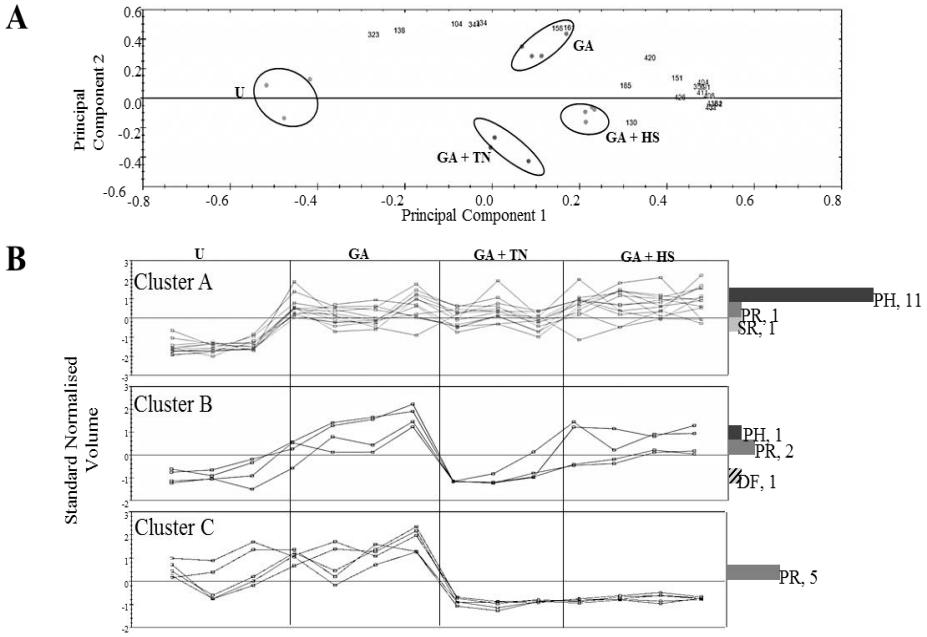


Figure 5. Principal Component Analysis (PCA) and clustering of extracellular protein spots. (A) PCA was performed on the 22 identified spots. Biological replicates are grouped by circles and spots are indicated by numbers. (B) Expression profiles of protein spots from PCA analysis were grouped in 3 clusters. Normalised volume for each spot is expressed relative to a reference gel (gibberellic acid), in which all spots volumes are by default set to 1. Corresponding functional classifications are indicated (PH: polysaccharide hydrolase; PR: protease; SR: stress response; and DF: defence). U: untreated; GA: gibberellic acid; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S2).

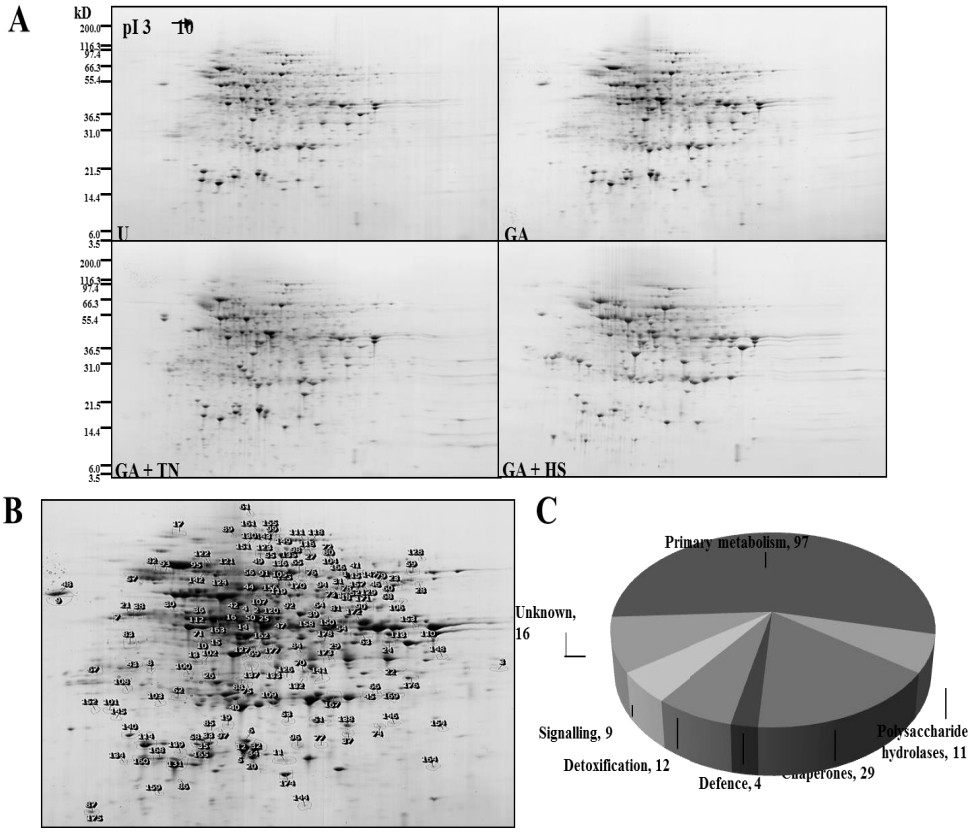


Figure 6. Intracellular water-soluble protein profiles of barley aleurone layer extracts**.** (A) One representative colloidal Coomassie blue-stained gel covering the pH range 3–10 is shown for each treatment. Molecular size markers are indicated. (B) Close-up view of the reference gel showing 178 identified spots. (C) Functional categories of the plant proteins identified. U: untreated; GA: gibberellic acid; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S3).

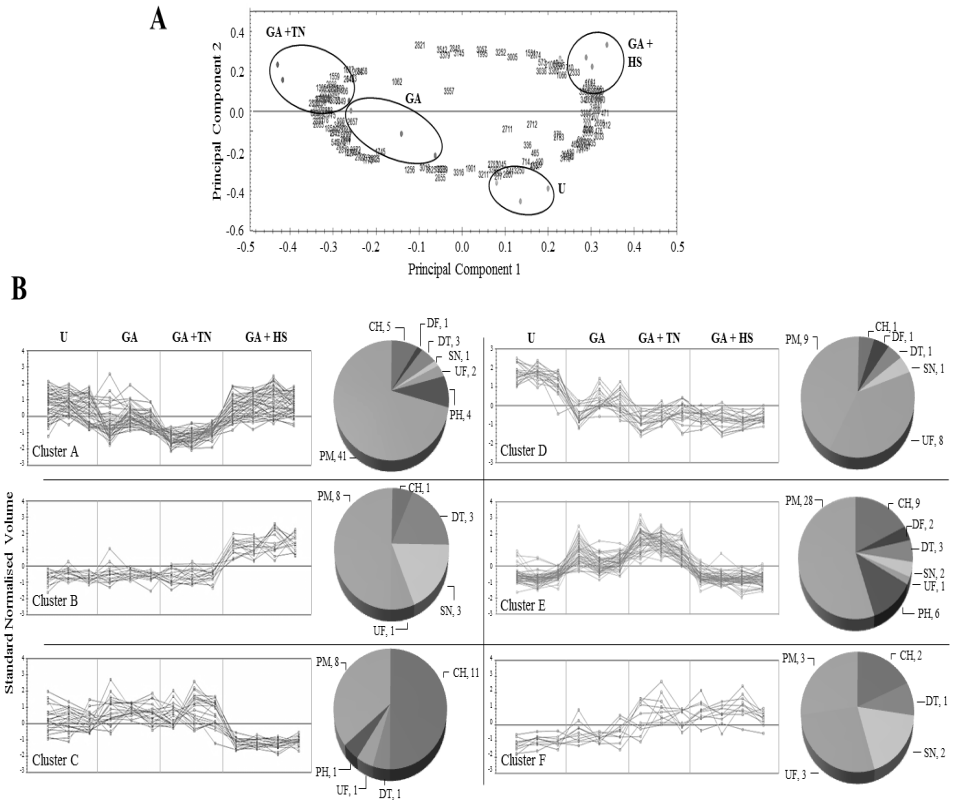


Figure 7. Principal Component Analysis (PCA) and clustering of intracellular water-soluble protein spots. (A) PCA was performed on the 178 identified spots. Biological replicates are grouped by circles and spots are indicated by numbers. (B) Expression profiles of protein spots from PCA analysis were grouped in 6 clusters. Normalised volume for each spot is expressed relative to a reference gel (gibberellic acid), in which all spots volumes are by default set to 1. Corresponding functional classifications are indicated (PM: primary metabolism; CH: chaperone; PH: polysaccharide hydrolase; DF: defence; DT: detoxification enzyme; SN: signalling; and UF: unknown function). U: untreated; GA: gibberellic acid; TN: tunicamycin; HS: heat shock. (For more details about protein identifications, see Supplemental Table S3).

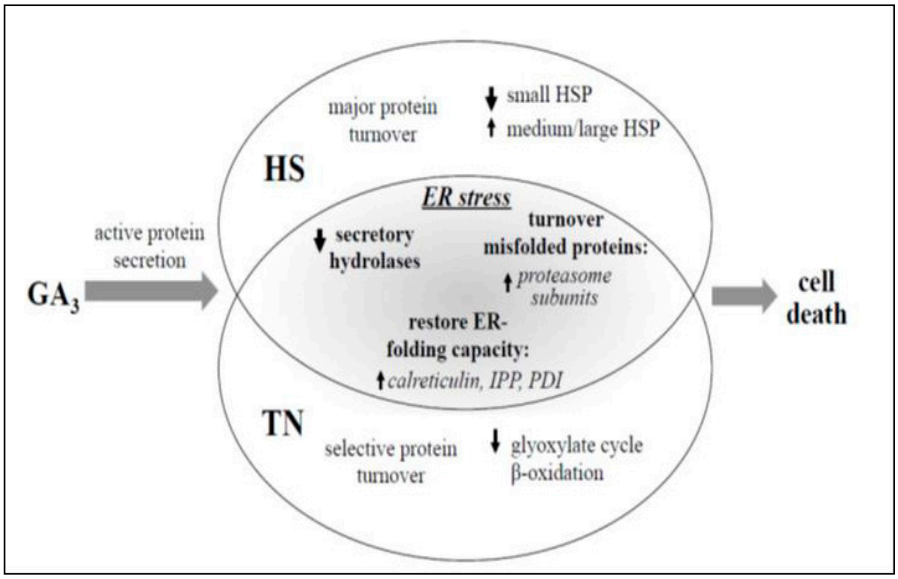


Figure 8. Summary scheme of the distinct and overlapping responses of gibberellic acid (GA)-induced aleurone layers to heat shock (HS) and tunicamycin (TN) treatments. 2D-base analysis suggest a common ER stress response, given by a reduction of protein secretion, a degradation of misfolded/unfolded proteins and a resumption of the ER-folding capacity by chaperones.

Chapter 5 — Conclusions and Remarks

Conclusions and Remarks

The aim of this project has mainly been to develop and apply proteomics techniques to map *N*-glycosylation sites in glycoproteins from cereal seeds. In chapter 2, a HILIC-based technique using a combination of ZIC-HILIC and cotton wool for glycopeptide enrichment (ZIC-cotton) was developed that allows for glycopeptide enrichment in a microcolumn format. This approach reduces co-enrichment of non-glycosylated peptides compared to using previously reported chromatography with cotton wool and 78 glycosylation sites in 66 different glycoproteins from the wheat albumin fraction were identified. In chapter 3, the same methodology was applied to barley aleurone layer proteins and 49 glycoproteins were identified.

The identification of these glycoproteins was based on protein sequence homology search against in silico-translated identified mRNA sequences present in TIGR (http://compbio.dfci.harvard.edu/tgi/) and NCBI (http://www.ncbi.nlm.nih.gov/) databases. For many of the identified mRNA sequences, no information is yet available about their encoded proteins, PTMs and functions. The present glycoprotein study contributes to validation of the presence of these proteins and their PTMs (*N*-glycosylations).

The supplement of ZIC-HILIC to cotton wool (ZIC-cotton) reduces the co-enrichment of low mass ranges (m/z 950-2500) non-glycosylated peptides, however the efficiency of ZIC-HILIC, cotton wool and ZIC-cotton for glycopeptide enrichment is difficult to evaluate. This is due to 1) a low number of glycopeptides derived from HRP-BSA used as a model because only a single glycoprotein (HRP) was used in the experiment 2) Several peaks obtained in the present study did not match known non-glycosylated or glycosylated peptides for example the peaks at m/z of 4174, 4012.5, 3370.9 and 2543.2. In an attempt to determine if these signals originate from glycopeptides, MS/MS spectra were acquired and examined for the presence of oxonium ions of *N*-acetylhexosamines (*m/z* 204) and *N*-acetylhexosamines after elimination of 1 or 2 mol of water, *m*/*z* 186 (HexNAc–H_2O) and 168 (HexNAc−$2H_2O$) (see appendix A). We detected oxonium ions of *m/z* 168, 186 and 204 from all known glycopeptides and many unknown peptides. However, HexNAc oxonium ions were also detected in some peptide precursors of non-glycosylated peptides used as negative controls. Therefore glycopeptide peaks cannot be assigned from this experiment. This might be because fragments of glycopeptides and non-glycosylated peptides are present at the same

mass range window and therefore a mixed population of fragmented ions from both non-glycosylated and glycosylated peptides were obtained.

Major food allergens are glycoproteins (Lack, 2008) and cross-reactions are very common for plant glycoproteins binding to patients IgE (Aalberse and van Ree, 1997, Mari et al., 1999). In the present study, potential allergenic glycoproteins from wheat and barley have been investigated by sequence homology search against proteins that have been proven to be allergenic. The findings show that many of the glycoproteins from wheat albumin (see chapter 2) and barley aleurone layer (see Chapter 3) have homology to known allergens and thereby provide new possible targets for further investigations on new allergens, allergenic cross-reactivity and protein glycan determinants.

In the present study, only a sub-glycoproteome of water-soluble albumins has been investigated. To complete whole wheat grain glycoproteome, further investigations on other wheat protein fractions e.g. wheat globulins, gliadins and glutenins (gliadin and glutenins together called as prolamins) should be performed. Besides the salt-soluble allergenic proteins discussed in chapter 2, several of the major water/salt-insoluble wheat flour proteins (prolamins) also appear to be implicated in baker's asthma (Laurière et al., 2006, Sandiford et al., 1997, Mittag et al., 2004).

We attempted to screen *N*-glycoproteins from wheat gliadins by the same procedure as performed in wheat albumins (chapter 1). However, a very low number of glycoproteins from the gliadin fraction were identified and they are most likely to be co-fractionated from the albumin fraction. This is probably because most of prolamin protein sequences are repetitive, rich in glutamine (Q) and proline (P), and tend to have few arginine (R) and lysine (K) residues required for trypsin digestion (Dupont et al., 2011). In addition, R/K followed by P cannot be a cleavage site by trypsin (Keil, 1992). This can result in too large peptides for detection by some MS instruments. The use of either less specific proteases such as pepsin, thermolysin and proteinase K (Chen et al., 2009) or the combination of several proteases for protein digestion may be applied to reduce this problem (Dupont et al., 2011, Chen et al., 2009).

PNGase A was chosen for deglycosylation in the present study because it cleaves all types of asparagine bound *N*-glycans including both with and without an α(1→3)-fucose. PNGase F cannot cleave *N*-glycans which are fucosylated at the α(1→3) position. However, the efficiency of PNGase A is much less than PNGase F. Thus, a combination of PNGase A and PNGase F may increase *N*-glycoproteome coverage and may give information about glycoproteins without α(1→3)-fucose that

might be important information for further development of e.g. therapeutic proteins produced recombinantly in plant.

In our analysis of wheat and barley proteomes we observed many peptides deamidated at asparagines (N) and glutamines (Q). For further analysis, it would be interesting to investigate if these deamidations occur due to chemical reactions in the experimental workflow or biological reactions in the cells. This information would improve the knowledge on post-translational modifications of proteins in cereals.

To provide further insight into protein structures and topology of membrane proteins in cereals, glycoprotein profiles and glycosylation site identifications are valuable information. In addition, these studies will be the basis for future development of a cereal glycoproteome database. Further studies on quantitative analysis of cereal glycoproteins will be very useful for comparative glycoproteome studies to reveal biological roles of glycoproteins and to develop biological protein markers in plants.

References

Aalberse, R. C.; van Ree, R. Crossreactive carbohydrate determinants. *Clin. Rev. Allergy Immunol.* **1997**, *15*, 375–387.

Chen R., Jiang X., Sun D., Han G., Wang F., Ye M., Wang L., Zou H. Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J. Proteome Res.* **2009**, *8*, 651–661.

Dupont, F. M.; Vensel, W. H.; Tanaka, C. K.; Hurkman, W. J.; Altenbach, S. B. Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Science.* **2011**, *9*, 1-10.

Keil. B. Specificity of Proteolysis. Springer-Verlag Berlin, Germany, **1992**, 66-69.

Lack, G. Food Allergy. *N. Engl. J. Med*. **2008**, *359*, 1252–1260.

Laurière, M.; Pecquet, C.; Bouchez-Mahiout, I.; Snégaroff, J.; Bayrou, O.; Raison-Peyron, N.; Vigan, M.; Hydrolysed wheat proteins present in cosmetics can induce immediate hypersensitivities. *Contact Dermatitis.* **2006**, *54,* 283-289.

Mari, A.; Iacovacci, P.; Afferni, C.; Barletta, B.; Tinghino, R.; Di Felice, G.; Pini, C. Specific IgE to cross-reactive carbohydrate determinants strongly affect the *in vitro* diagnosis of allergic diseases. *J. Allergy Clin. Immunol.* **1999**, *103*, 1005–1011.

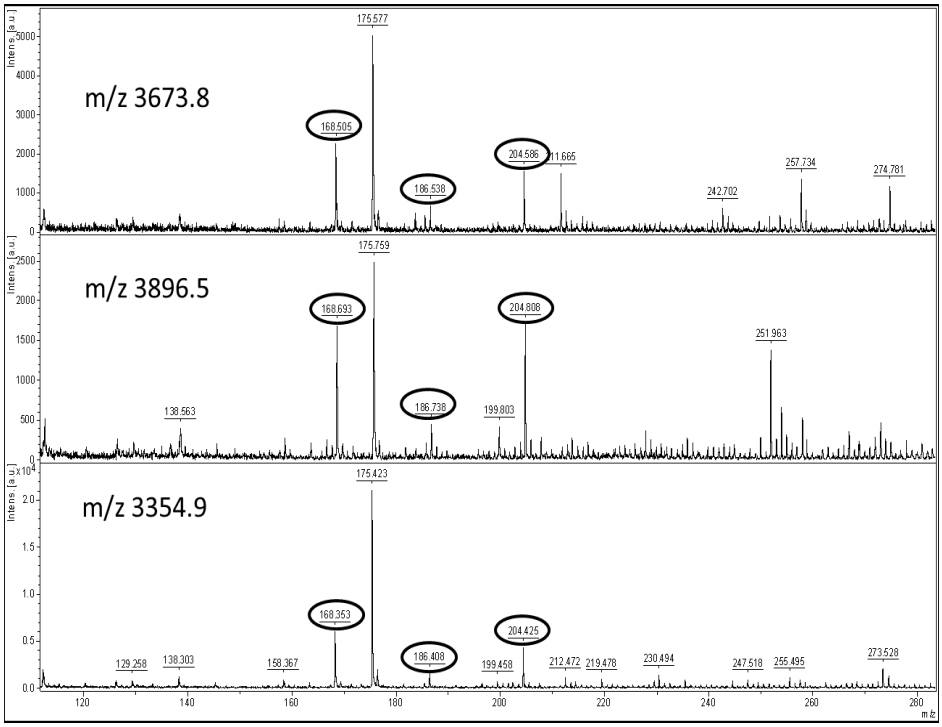
Mittag, D.; Niggermann, B.; Sander, I.; Reese, I.; Fiedler, E. M.; Worm, M.; Vieths, S.; Reese, G. Immunoglobulin E-reactivity of wheat-allergic subjects (baker's asthma, food allergy, wheat dependent, exercise-induced anaphylaxis) to wheat protein fractions with different solubility and digestibility. *Mol. Nutr. Food Res.* **2004**, *48*, 380-389.

Sandiford, C. P.; Tatham, A. S.; Fido, R.; Welch, J. A.; Jones, M. G.; Tee, R. D.; Shewry, P. R.; Newman-Taylor, A. J. Identification of the major water/salt insoluble wheat proteins involved in cereal hypersensitivity. *Clin.Exp. Allergy.* **1997**, *27*, 1120-1129.

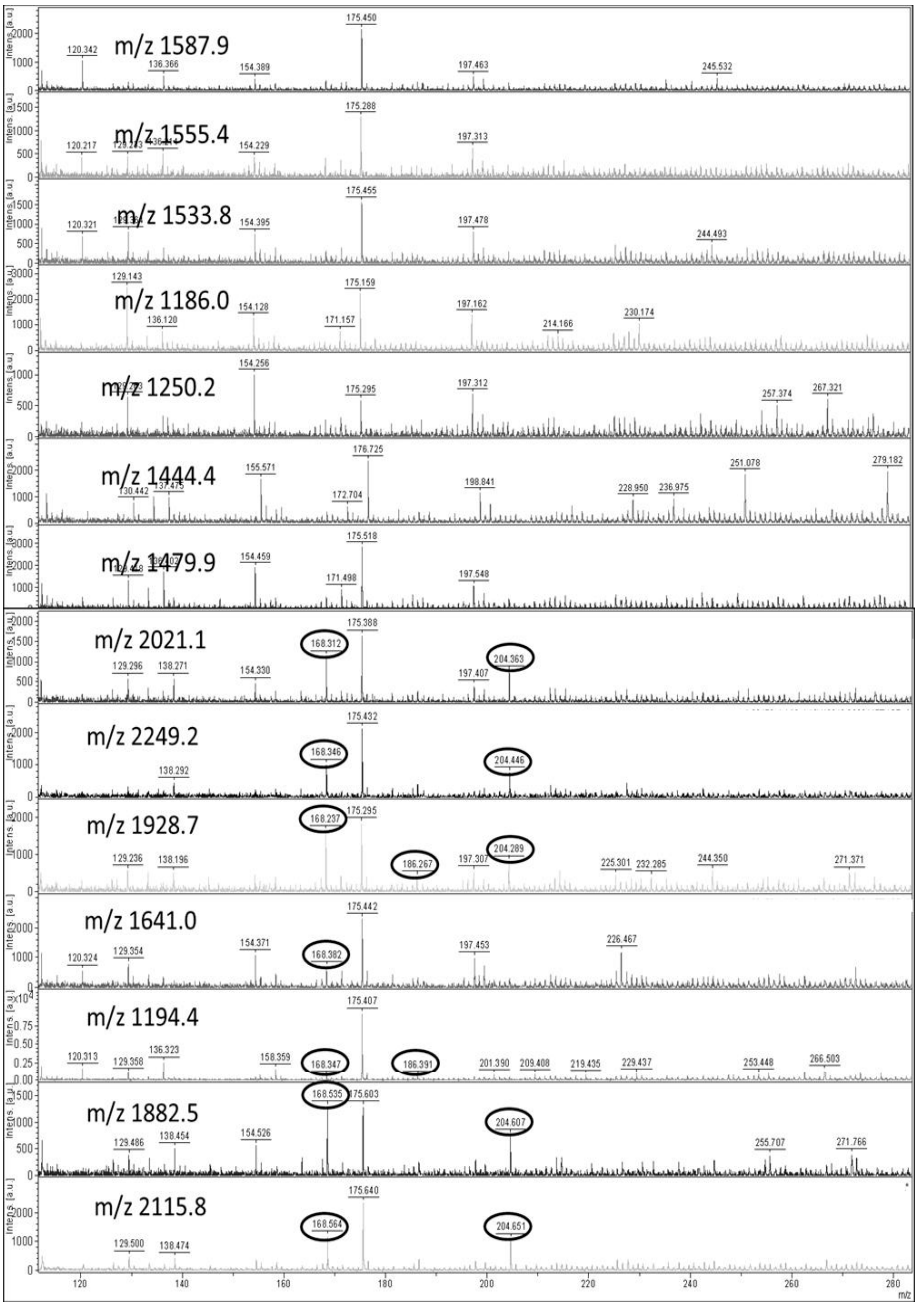
Precursor ion scanning for oligosaccharide derived fragment ions to detect all possible HRP-BSA glycopeptides.

In order to detect all possible HRP-BSA glycopeptides precursor ions from the samples, precursor ion scanning for oligosaccharide derived fragment ions of N-acetylhexosamines at m/z 204 and N-acetylhexosamines after elimination of 1 or 2 mol of water, m/z 186 (HexNAc–H_2O) and 168 (HexNAc−$2H_2O$), are used. The MS/MS fragmentations of precursor ions were performed on several HRP-BSA peptides matched to previously reported glycosylated peptides (A) used as positive control, peptides matched to predicted non-glycosylated peptides (B) used as negative control and unknown peptides (C). In this experiment, oxonium ions m/z 168, 186 and 204 were detected from all known glycopeptides and some unknown peptides. However, the oxonium ions could also be detected in some peptide precursors of non-glycosylated peptides. Therefore the unknown glycopeptide peaks cannot be determined if they are glycosylated in this experiment.

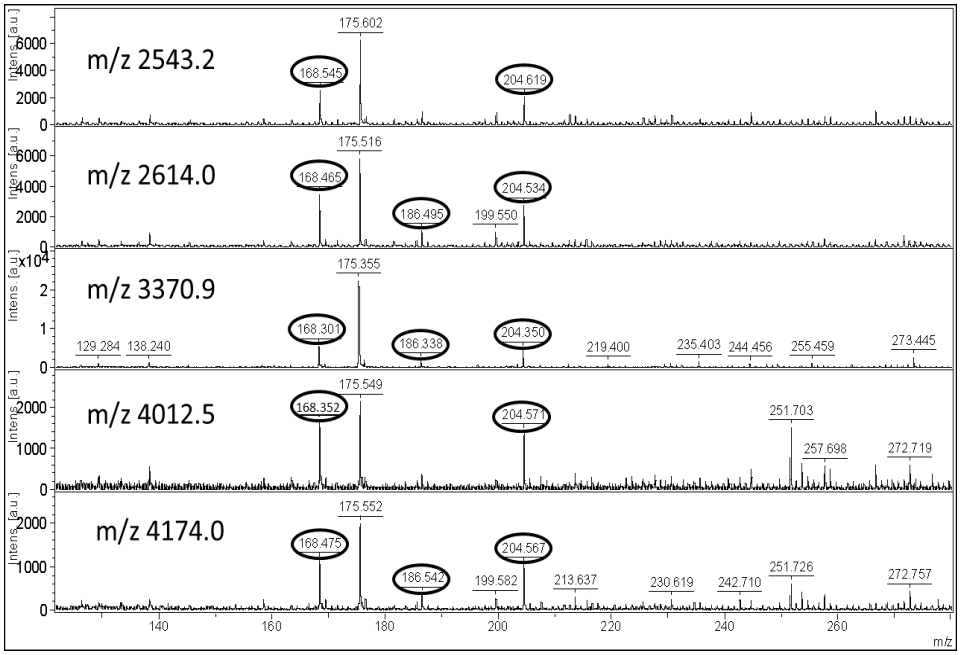
A. Peptides with m/z matched to predicted glycosylated peptides



B. m/z matched to predicted non glycosylated peptides



C. Unknown peptides



Enzyme and Protein Chemistry (EPC) works within protein biochemistry, carbohydrate biochemistry, molecular biology, microbiology and plant biochemistry. The main activities of EPC are related to food and raw materials for food, but the methods and main strategies are relevant to various biotechnological issues. The aim is to explain the molecular mechanisms and interactions relevant for functionality and quality of foodstuffs and raw materials and to identify the biochemical mechanisms connected to health promoting and nutrition related effects of food.

Enzyme and Protein Chemistry
Department of Systems Biology
Technical University of Denmark
Søltofts Plads, Building 224
DK-2800 Kgs. Lyngby
Denmark

Phone: +45 4525 2525
Fax: +45 4588 6307

Web: www.bio.dtu.dk

ISBN : 978-87-91494-76-5